

Weak limits for exploratory plots in the analysis of extremes

Bikramjit Das and Souvik Ghosh

*RiskLab, Department of Mathematics,
ETH Zurich,
Rämistrasse 101,
8092 Zurich, Switzerland
e-mail: bikram@math.ethz.ch*

*Department of Statistics,
Columbia University,
1255 Amsterdam Avenue,
New York, NY 10027, USA
e-mail: ghosh@stat.columbia.edu*

Abstract: Exploratory plotting tools have been devised aplenty in order to diagnose the goodness-of-fit of data sets to a hypothesized distribution. Some of them have found extensive use in diverse areas of finance, telecommunication, environmental science, etc. in order to detect sub-exponential or heavy-tailed behavior in observed data. In this paper we concentrate on two such plotting methodologies: the Quantile-Quantile plots for heavy-tails and the Mean Excess plots. Under the assumption of heavy-tailed behavior of the underlying sample the convergence in probability of these plots to a fixed set in a suitable topology of closed sets of \mathbb{R}^2 has been studied in [Das and Resnick \(2008\)](#) and [Ghosh and Resnick \(2010\)](#). These results give theoretical justifications for using the plots to test the null hypothesis that the underlying distribution is heavy-tailed by checking if the observed plot is “close” to the limit under the null hypothesis. In practice though one set of observations would lead to only one plot of each kind. This is often not good enough to verify proximity of the plot to the fixed set of interest under the null hypothesis of heavy-tails. An appropriate confidence bound around the plots covering or not covering the fixed set would stand as a better indicator for such a purpose. In this paper we provide weak limits for Quantile-Quantile plots for heavy-tails and Mean Excess plots when the underlying distribution of the sample is heavy-tailed. As an application we are able to construct confidence bounds around the plots which enables us to check whether the underlying distribution is heavy-tailed or not.

AMS 2000 subject classifications: Primary 62G32, 60G70; secondary 62G10, 62G15, 60F05.

Keywords and phrases: regular variation, random sets, QQ plots, ME plots, extreme values, weak limits, confidence bounds.

1. Introduction

Statistical analysis of extremes in available data has been quite important in a variety of areas like finance ([McNeil et al., 2005](#)), telecommunication ([Maulik et al., 2002](#); [D’Auria and Resnick, 2006](#)), hydrology ([Katz et al., 2002](#)), environmental statistics ([Davison and Smith, 1990](#); [Smith, 2003](#)) and many more. Prior to analyzing features of the data using extreme value analysis it is imperative that we check whether extreme-value modeling is well-suited in the given context. Popular exploratory techniques in this direction have been the Mean Excess (ME) plots ([Davison and Smith, 1990](#)) and the Quantile-Quantile (QQ) plots which are specifically tuned for heavy-tailed data ([Kratz and Resnick, 1996](#)). These plots are used to detect whether or not the data is heavy-tailed in the sense that the tails of the underlying distributions are regularly varying with some tail index $-1/\xi$ where $\xi > 0$ ([Resnick, 2008](#), Chapter 1). It has been shown earlier that under this assumption of regular variation of tails both the QQ plot for heavy-tails and the ME plot when suitably normalized converge in probability to a fixed closed set, which is a straight line through the origin in most

*The authors are thankful to Paul Embrechts (ETH Zurich), Sidney I. Resnick (Cornell University) and Gennady Samorodnitsky (Cornell University) for their detailed comments on a draft of the paper which greatly helped in improving the paper. Souvik Ghosh was partially supported by the FRAP program at Columbia University.

cases (i.e., except for the ME plot for data where the mean does not exist, in which case the limit is neither a straight line nor a fixed set); see [Das and Resnick \(2008\)](#) and [Ghosh and Resnick \(2010\)](#). These results give proper theoretical justification for the use of the two plots mentioned to test the hypothesis that the data comes from a heavy-tailed distribution. But it should be noted that one data set only leads to one plot of each kind. Clearly one plot is not enough to detect the proximity between the plots as a closed set and the intended fixed set under the assumption of heavy-tails. This is clarified with examples in [Section 6](#). One way to get around this will be to provide confidence bounds around these plots and see if these bounds contain the fixed set of interest. It is also possible to bootstrap more data sets from the one that we have and superimpose these plots and see if it covers the fixed set of interest. The bootstrapping method is often used in practice but bootstrapping in a heavy-tailed set up needs additional care while using them, for details on the difficulties with this see ([Resnick, 2007](#), Chapter 6.4). In our approach we take the first route of finding confidence bounds for these plots. First we study weak limits of the plots for heavy-tailed data and then using these limits we describe methods of obtaining confidence bounds for both the QQ plots for heavy-tails and the ME plots.

1.1. Plan for this paper

We start by introducing the QQ plot and the ME plot in [Section 1](#) and by explaining the problem we are looking at. The plots that we talk about in this paper can all be visualized as random closed sets in \mathbb{R}^2 , hence notions of convergence also must be described in an appropriate sense here. In [Section 2](#) we set up necessary tools to talk about convergence of random closed sets in \mathbb{R}^2 . Also note that both the QQ plot and the ME plot depend on the tail empirical measure of the sample, that is, both the plots depend on a portion of the data that are made out of the upper order statistics of the sample. So we will provide distributional convergence results for the plots under the null hypothesis of the tails of the underlying distributions being regularly varying using convergences of appropriate functionals of the tail empirical measure. This is discussed in detail in [Sections 3](#) and [4](#). Now, as an application to the weak limits we obtained for these plots under the null hypothesis, we construct confidence bounds for them in [Section 5](#). Finally, in [Section 6](#) we use the results obtained in the previous sections to construct confidence bounds on simulated and real data sets to exemplify how they perform in practice. We conclude in [Section 7](#) along with a discussion on future directions.

1.2. QQ plots for heavy-tails

Suppose we want to test the null hypothesis that observations from a sample are i.i.d. from some known distribution F . One quite intuitive way to go is to make a QQ plot which is a plot of the empirical quantiles from the data against the distributional quantiles of F . The QQ plot has been popular for graphically detecting the goodness-of-fit for a sample to the distribution F . If the true distribution of the sample is F then the QQ plot should converge, in an appropriate sense, to a straight line. Results involving empirical process and quantile process convergences are available in [Shorack and Wellner \(1986\)](#) which can be appropriately used to create confidence intervals for QQ plots. The QQ plot we consider is a little different and specifically designed to check for distributions F where $1 - F$ is regularly varying with some index $-1/\xi$, $\xi > 0$, also denoted $1 - F \in RV_{-1/\xi}$ ([Resnick, 2008](#), Chapter 1). For a sample X_1, X_2, \dots, X_n , its decreasing order statistics is denoted by $X_{(1)} \geq X_{(2)} \geq \dots \geq X_{(n)}$ and the QQ plot in this context is defined by

$$\mathcal{Q}_n = \left\{ \left(-\log \frac{j}{k}, \log \frac{X_{(j)}}{X_{(k)}} \right) : 1 \leq j \leq k \right\}, \quad k < n.$$

Clearly we concentrate on the top k quantiles of the data justified by the fact that $1 - F \in RV_{-1/\xi}$ only provides us with information about the right tail of the data. Under the null hypothesis of $1 - F \in RV_{-1/\xi}$ for some $\xi > 0$, [Das and Resnick \(2008\)](#) has shown convergence in probability for QQ plots in an appropriate topology of random closed sets when the data is assumed to be an independent and identically distributed (i.i.d.) sample.

1.3. ME Plots

The ME function of a random variable X is defined as:

$$M(u) := E[X - u | X > u], \quad (1.1)$$

provided $EX_+ < \infty$, and is also known as the *mean residual life function*. A natural estimate of $M(u)$ is the empirical ME function $\hat{M}(u)$ defined as

$$\hat{M}(u) = \frac{\sum_{i=1}^n (X_i - u) I_{[X_i > u]}}{\sum_{i=1}^n I_{[X_i > u]}}, \quad u \geq 0. \quad (1.2)$$

The ME plot is the plot of the points $\{(X_{(k)}, \hat{M}(X_{(k)})) : 1 < k \leq n\}$.

The ME plot is often used as a simple graphical test to check if data conform to a generalized Pareto distribution (GPD). The GPD is an important class of distributions and is fundamental for the peaks over threshold method. The GPD is characterized by its cumulative distribution function $G_{\xi, \beta}$:

$$G_{\xi, \beta}(x) = \begin{cases} 1 - (1 + \xi x / \beta)^{-1/\xi} & \text{if } \xi \neq 0 \\ 1 - \exp(-x/\beta) & \text{if } \xi = 0 \end{cases} \quad (1.3)$$

where $\beta > 0$, and $x \geq 0$ when $\xi \geq 0$ and $0 \leq x \leq -\beta/\xi$ if $\xi < 0$. The parameters ξ and β are referred to as the *shape* and *scale* parameters respectively. For a Pareto distribution, the tail index α is just the reciprocal of ξ when $\xi > 0$.

For a random variable $X \sim G_{\xi, \beta}$, we have $E(X) < \infty$ iff $\xi < 1$ and in this case, the ME function of X is linear in u :

$$M(u) = \frac{\beta}{1 - \xi} + \frac{\xi}{1 - \xi} u, \quad (1.4)$$

where $0 \leq u < \infty$ if $0 \leq \xi < 1$ and $0 \leq u \leq -\beta/\xi$ if $\xi < 0$. In fact, the linearity of the ME function characterizes the GPD class, cf. McNeil et al. (2005); Embrechts et al. (1997). Davison and Smith (1990) used this property and suggested that if the ME plot is close to linear for high values of the threshold then there is no evidence against the use of a GPD model. See also Embrechts et al. (1997) and Hogg and Klugman (1984) for the implementation of this plot in practice. Ghosh and Resnick (2010) discusses convergence in probability for the high thresholds of suitably normalized ME plots in an appropriate topology of random closed sets when the data is an i.i.d. sample.

The advantage of the ME plot over the QQ plot is that it works when $-\infty < \xi < 1$, i.e., when the sample is in the maximal domain of attraction of the Weibull or the Gumbel. In this paper though we restrict to the case when $\xi > 0$ which is the case of maximal domain of attraction of the Fréchet distribution. The disadvantage of the ME plot is that it requires $\xi < 1$ to make proper sense of the result, in other words, for the ME plot to make sense the underlying distribution should have a finite mean, although limits can and has been obtained for the ME plots even when the distributional mean is not finite.

2. Preliminaries

2.1. Topology on closed sets of \mathbb{R}^2

Since we are dealing with plots which are closed sets in \mathbb{R}^2 it is imperative to understand the topology on closed sets. We denote the collection of all closed (compact) sets in \mathbb{R}^2 by \mathcal{F} (\mathcal{K} respectively). We consider a hit and miss topology on \mathcal{F} called the Fell topology. The Fell topology is generated by the families $\{\mathcal{F}^K, K \text{ compact}\}$ and $\{\mathcal{F}_G, G \text{ open}\}$ where for any set B

$$\mathcal{F}^B = \{F \in \mathcal{F} : F \cap B = \emptyset\} \quad \text{and} \quad \mathcal{F}_B = \{F \in \mathcal{F} : F \cap B \neq \emptyset\}$$

Hence \mathcal{F}^B and \mathcal{F}_B are collections of closed sets which miss and hit the set B , respectively. This is the reason for which such topologies are called hit and miss topologies. In the Fell topology a sequence of closed sets $\{F_n\}$ converges to $F \in \mathcal{F}$ if and only if the following two conditions hold:

- F hits an open set G implies there exists $N \geq 1$ such that for all $n \geq N$, F_n hits G .
- F misses a compact set K implies there exists $N \geq 1$ such that for all $n \geq N$, F_n misses K .

The Fell topology on the closed sets of \mathbb{R}^2 is metrizable and we indicate convergence in this topology of a sequence (F_n) of closed sets to a limit closed set F by $F_n \rightarrow F$. Often though, it is easier to deal with the following characterization of convergence.

Lemma 2.1. *A sequence $F_n \in \mathcal{F}$ converges to $F \in \mathcal{F}$ in the Fell topology if and only if the following two conditions hold:*

1. For any $t \in F$ there exists $t_n \in F_n$ such that $t_n \rightarrow t$.
2. If for some subsequence (m_n) , $t_{m_n} \in F_{m_n}$ converges, then $\lim_{n \rightarrow \infty} t_{m_n} \in F$.

See Theorem 1-2-2 in (Matheron, 1975, p.6) for a proof of this Lemma.

Let $\sigma_{\mathcal{F}}$ denote the Borel σ -algebra generated by the Fell topology of open sets (not to be confused with open sets in \mathbb{R}^d). A random closed set $X : \Omega \mapsto \mathcal{F}$ is a measurable mapping from $(\Omega, \mathcal{A}, P')$ to $(\mathcal{F}, \sigma_{\mathcal{F}})$. Denote by P the induced probability on $\sigma_{\mathcal{F}}$, that is, $P = P' \circ X^{-1}$.

Since the Fell topology is metrizable, the definition of convergence in probability is obvious. The following result is a well-known and helpful characterization for convergence in probability of random variables and it holds for random sets as well; see Theorem 6.21 in (Molchanov, 2005, p.92).

Lemma 2.2. *A sequence of random sets (F_n) in \mathcal{F} converges in probability to a random set F if and only if for every subsequence (n') of \mathbb{Z}_+ there exists a further subsequence (n'') of (n') such that $F_{n''} \rightarrow F$ -a.s.*

A sequence of random closed sets $(X_n)_{n \geq 1}$ weakly converges to a random closed set X with distribution P if the corresponding induced probability measures $(P_n)_{n \geq 1}$ converge weakly to P , i.e.,

$$P_n(\mathcal{B}) = P'_n \circ X_n^{-1}(\mathcal{B}) \rightarrow P(\mathcal{B}) = P' \circ X^{-1}(\mathcal{B}), \quad \text{as } n \rightarrow \infty,$$

for each $\mathcal{B} \in \sigma_{\mathcal{F}}$ such that $P(\partial \mathcal{B}) = 0$. This is not always straightforward to verify from the definition. The following characterization of weak convergence in terms of sup-measures is very useful, cf. (Vervaat, 1997). Suppose $h : \mathbb{R}^d \rightarrow \mathbb{R}_+ = [0, \infty)$. For $X \subset \mathbb{R}^d$, define $h(X) = \{h(x) : x \in X\}$ and h^\vee is the sup-measure generated by h defined by

$$h^\vee(X) = \sup\{h(x) : x \in X\}$$

(Molchanov, 2005; Vervaat, 1997). These definitions permit the following characterization (Molchanov, 2005, page 87).

Lemma 2.3. *A sequence $(X_n)_{n \geq 1}$ of random closed sets converges weakly to a random closed set X if and only if $\mathbb{E}h^\vee(X_n)$ converges to $\mathbb{E}h^\vee(X)$ for every non-negative continuous function $h : \mathbb{R}^d \rightarrow \mathbb{R}$ with a bounded support.*

We use the following notation: For a real number x and a set $A \subset \mathbb{R}^n$, $xA = \{xy : y \in A\}$ and $x + A = \{x + y : y \in A\}$.

Matheron (1975) and Molchanov (2005) are good references for the theory of random sets.

2.2. Miscellany

Throughout this paper we will take $k := k_n$ to be a sequence increasing to infinity such that $k_n/n \rightarrow 0$. For a distribution function $F(x)$ we write $\bar{F}(x) := 1 - F(x)$ for the tail and the quantile function is

$$b(u) := F^{\leftarrow}(1 - \frac{1}{u}) = \inf \left\{ s : F(s) \geq 1 - \frac{1}{u} \right\} = \left(\frac{1}{1 - F} \right)^{\leftarrow}(u).$$

A function $U : (0, \infty) \rightarrow \mathbb{R}_+$ is regularly varying with index $\rho \in \mathbb{R}$, written $U \in RV_\rho$, if

$$\lim_{t \rightarrow \infty} \frac{U(tx)}{U(t)} = x^\rho, \quad x > 0.$$

It is discussed in several books such as Resnick (2007, 2008); Seneta (1976); Geluk and de Haan (1987); de Haan (1970); de Haan and Ferreira (2006); Bingham et al. (1987).

We use $M_+(0, \infty]$ to denote the space of nonnegative Radon measures μ on $(0, \infty]$ metrized by the vague metric. Point measures are written as a function of their points $\{x_i, i = 1, \dots, n\}$ by $\sum_{i=1}^n \delta_{x_i}$. See, for example, (Resnick, 2008, Chapter 3).

We will use the following notations to denote different classes of functions: For $0 \leq a < b \leq \infty$,

1. $\mathbb{C}[a, b)$: Continuous functions on $[a, b)$.
2. $\mathbb{D}[a, b)$: Right-continuous functions with finite left limits defined on $[a, b)$.
3. $\mathbb{D}_l[a, b)$: Left-continuous functions with finite right limits defined on $[a, b)$.

$\mathbb{D}[0, 1]$ is complete and separable under a metric $d_0(\cdot)$ which is equivalent to the Skorohod metric $d_S(\cdot)$ (Billingsley, 1968, page 128), but not under the uniform metric $\|\cdot\|$. As we will see later, the limit processes that appear in our analysis below are always continuous. We can check that if x is continuous (in fact uniformly continuous) in $[0, 1]$, for $x_n \in D[0, 1]$, $\|x_n - x\| \rightarrow 0$ is equivalent to $d_S(x_n, x) \rightarrow 0$ and hence equivalent to $d_0(x_n, x) \rightarrow 0$ as $n \rightarrow \infty$ (Billingsley, 1968, page 124). So we use convergence in uniform metric, for our convenience henceforth. For spaces of the form $\mathbb{D}[a, b)$ or $\mathbb{D}_l[a, b)$ we will consider the topology of locally uniform convergence. In some cases we will also consider product spaces of functions and then the topology will be the product topology. For example, $D_l^2[1, \infty)$ will denote the class of 2-dimensional functions on $[1, \infty)$. The classes of functions defined on the sets $[a, b]$ or $(a, b]$ will have the obvious notation.

3. Limit results for the QQ plots

Convergence of empirical processes and quantile processes to functionals of Gaussian processes, usually Brownian motion and Brownian bridges are quite well-known, cf. Shorack and Wellner (1986). We prove similar results for extreme order statistics. We use the weak limit of tail empirical measure and deduce weak convergence of the logarithmic version of the QQ plot of the extreme order statistics as a random set.

First we discuss some limit results in the context of QQ plots that are already known. The following was proved is Das and Resnick (2008):

Proposition 3.1. *Suppose X_1, \dots, X_n are i.i.d. with common distribution F and $X_{(1)} \geq X_{(2)} \geq \dots \geq X_{(n)}$ are the order statistics from this sample. If F is strictly increasing and continuous on its support, then*

$$\mathcal{T}_n := \left\{ \left(F^{\leftarrow} \left(\frac{i}{n+1} \right), X_{(n-i+1)} \right); 1 \leq i \leq n \right\} \xrightarrow{P} \mathcal{T} := \left\{ (x, x) : x \in \text{support}(F) \right\}$$

in \mathcal{F} .

The proposition was first proved for the case where F is the uniform distribution on $[0, 1]$ and then it was extended to general distributions on \mathbb{R} . This proposition though is not enough if one is interested in creating confidence bounds from the data. For that purpose one would need weak convergence results. These results are widely known when the distribution F is known. We briefly state the results below but would refer to (Shorack and Wellner, 1986, Chapter 3) for further details.

3.1. QQ plots for a sample with known target distribution F

We start with the simplest case. Suppose U_1, U_2, \dots, U_n are i.i.d. $U(0, 1)$. Denote the order statistics of this sample by $U_{(1)} \geq U_{(2)} \geq \dots \geq U_{(n)}$. then

$$\mathcal{T}_n = \left\{ \left(\frac{i}{n+1}, U_{(n-i+1)} \right), 1 \leq i \leq n \right\} \quad \text{and} \quad \mathcal{T} = \left\{ (x, x) : 0 \leq x \leq 1 \right\}. \quad (3.1)$$

Then Proposition 3.1 yields $\mathcal{T}_n \xrightarrow{a.s.} \mathcal{T}$ in \mathcal{K} . Now consider the empirical process

$$F_n(t) = \frac{1}{n} \sum_{i=1}^n I_{[U_i \leq t]} \quad \text{for } 0 \leq t \leq 1.$$

and the quantile process defined as

$$\begin{aligned} G_n(t) &:= F_n^{\leftarrow}(t) \equiv \inf\{x : F_n(x) \geq t\} \\ &= U_{(n-i+1)} \quad \text{for } \frac{i-1}{n} < t \leq \frac{i}{n} \text{ and } 1 \leq i \leq n \end{aligned}$$

and $G_n(0) = 0$. Let $\mathcal{T}_n^* = \{(t, G_n(t)) : 0 \leq t \leq 1\}$. Clearly $\mathcal{T}_n \subset \mathcal{T}_n^*$ in \mathbb{R}^2 . We can also check that $\mathcal{T}_n \xrightarrow{a.s.} \mathcal{T}$ in \mathcal{F} is equivalent to $\mathcal{T}_n^* \xrightarrow{a.s.} \mathcal{T}^*$ in \mathcal{F} .

Now keeping weak limit in mind, define

$$B_n(t) = \sqrt{n}(G_n(t) - t).$$

Let B denote a Brownian Bridge on $[0, 1]$, i.e., a Gaussian process on $\mathbb{C}[0, 1]$ with

$$E(B(t)) = 0 \quad \text{and} \quad \text{cov}(B(s), B(t)) = s \wedge t - st \quad \forall 0 \leq s, t \leq 1.$$

It is a simple extension of the CLT to check that $B_n \Rightarrow B$ in finite dimensional distributional sense. In fact, we can define the uniform random variables U_i and B on the same probability space $(\Omega, \mathcal{A}, \mathbf{P})$ such that $\|B_n - B\| \rightarrow 0$ almost surely (Shorack and Wellner, 1986, Chapter 3).

This result for the uniform case can be easily extended to the case where X_1, X_2, \dots, X_n are i.i.d. from some distribution F which has density $f = F'$ and also $g = (F^{-1})'$. As in the uniform case we define:

$$\begin{aligned} G_n(t) &:= F_n^{\leftarrow}(t) \equiv \inf\{x : F_n(x) \geq t\} \\ &= X_{(n-i+1)} \quad \text{for } \frac{i-1}{n} < t \leq \frac{i}{n} \text{ and } 1 \leq i \leq n. \end{aligned}$$

and $G_n(0) = 0$. Now define

$$B_n(t) = \sqrt{n}g(t)(G_n(t) - G(t)).$$

where $G(t) := F^{\leftarrow}(t)$, $0 \leq t \leq 1$. Here again we know that $B_n \rightarrow B$ in fidi's and an almost sure convergence is possible under a special construction (Shorack and Wellner, 1986).

3.2. QQ plots in the regularly varying case

Now assume that X_1, X_2, \dots, X_n are i.i.d. from a distribution F . Suppose we want to check from the sample whether F is heavy-tailed or not, more specifically, we want to check whether $1 - F$ is regularly varying with some index $-1/\xi$, i.e., $\bar{F} \in RV_{-1/\xi}$, with $\xi > 0$. Hence in the sense of testing a hypothesis, our null hypothesis is that $\bar{F} \in RV_{-1/\xi}$ for some $\xi > 0$, but we really do not have any specific form for F . We define the following sets:

$$\mathcal{Q}_n = \left\{ \left(-\log \frac{j}{k}, \log \frac{X_{(j)}}{X_{(k)}} \right) : 1 \leq j \leq k \right\}, \quad k < n, \quad (3.2)$$

$$\mathcal{Q} = \left\{ (x, \xi x) : x \geq 0 \right\}. \quad (3.3)$$

The set \mathcal{Q}_n is the logarithmic version of the QQ plot for the first k order statistics from the sample X_1, \dots, X_n .

We know that under the null hypothesis, $\mathcal{Q}_n \xrightarrow{P} \mathcal{Q}$ as $k, n \rightarrow \infty$ with $k/n \rightarrow 0$ from Das and Resnick (2008). We show henceforth that a distributional convergence can also be obtained in this case.

Assumption 3.2. We say that F satisfies condition R if

$$\lim_{n \rightarrow \infty} \sqrt{k} \left(\frac{n}{k} \bar{F}(b(n/k)y^{-\xi}) - y \right) = 0 \quad (3.4)$$

locally uniformly on $(0, \infty]$ as $k, n, n/k \rightarrow \infty$.

Theorem 3.3. Suppose X_1, X_2, \dots, X_n are i.i.d. F , where $\bar{F} \in RV_{-1/\xi}$, with $\xi > 0$ and satisfies condition R . Then

$$\begin{aligned} \mathcal{QN}_n &:= \left\{ \left(-\log \frac{j}{k}, -\xi \log \frac{j}{k} + \sqrt{k} \left(\log \frac{X_{(j)}}{X_{(k)}} + \xi \log \frac{j}{k} \right) \right) : 1 \leq j \leq k \right\} \\ &\Rightarrow \mathcal{QN} := \left\{ \left(-\log t, -\xi \log t + \xi t^{-1} B(t) \right) : 0 < t \leq 1 \right\} \quad \text{in } \mathcal{F}, \end{aligned}$$

where $B(t)$ is a Brownian Bridge on $[0, 1]$ restricted to $(0, 1]$.

Remark 3.4. The set \mathcal{QN}_n is a suitably normalized version of the QQ plot which allows us to obtain a weak limit. It is important to observe that the format in which we have expressed the result is not standard in the literature as far as weak limits of random variables or functions are concerned. Usual weak limit results will only consider the normalized difference of the random variable from its mean or its limit in probability. In our setting it is imperative to state the result in the form which we have used. We look at the plot as the probability limit perturbed by the normalized deviation around it; i.e., we shift the normalized differences so that we can obtain the distribution of the deviation of the observed points of the QQ plot from its mean position. If we do not make this shift, the weak limit will always hover around the y -axis and will not give the deviation from the actual point in the plot.

Remark 3.5. We have assumed Condition R in order to prove the a weak limit for the QQ plots. Without this assumption we can show the convergence of tail empirical measure with unknown centering $\frac{n}{k} \bar{F}(b(n/k)y^{-\xi})$ as in (3.7) but we wish the centering to be y here. To achieve this

$$\lim_{n \rightarrow \infty} \sqrt{k} \left(\frac{n}{k} \bar{F}(b(n/k)y^{-\xi}) - y \right)$$

should exist and have a finite limit which we assume to be 0 without loss of any generality. The same theorem can be proved by replacing Condition R with the stronger condition of *second order regular variation* [de Haan and Ferreira \(2006\)](#); [de Haan and Stadtmueller \(1996\)](#); [de Haan and Peng \(1998\)](#). Neither condition R nor the second order RV condition are easy to check in practice, albeit we resort to assuming them in order to obtain distributional limits.

Proof. We will derive this result from the weak convergence of the tail empirical measure. The tail empirical measure is defined as

$$\nu_n(\cdot) := \frac{1}{k} \sum_{i=1}^n \epsilon_{X_i/b(n/k)}(\cdot) \quad (3.5)$$

as a random element of $M_+(0, \infty]$ where $\epsilon_x(\cdot)$ puts unit mass at x . By Theorem 4.1 ([Resnick, 2007](#), p.79) we get that

$$\nu_n \Rightarrow \nu \quad \text{in } M_+(0, \infty] \quad (3.6)$$

where $\nu(y, \infty] = y^{-1/\xi}$, $y > 0$. Furthermore, Theorem 9.1 in ([Resnick, 2007](#), p.292) gives us

$$\sqrt{k} \left(\nu_n(y^{-\xi}, \infty] - \frac{n}{k} \bar{F}(b(n/k)y^{-\xi}) \right) \Rightarrow W(y) \quad \text{in } D[0, \infty). \quad (3.7)$$

where W is a standard Brownian motion on $[0, \infty)$. Since F satisfies condition R we obtain

$$\sqrt{k} \left(\nu_n(y^{-\xi}, \infty] - y \right) \Rightarrow W(y) \quad \text{in } D[0, \infty). \quad (3.8)$$

We will use this to find the limiting distribution of

$$\sqrt{k} \left(\log \frac{X_{(\lceil kt \rceil)}}{X_{(k)}} + \xi \log t \right) = \sqrt{k} \log \left(\frac{X_{\lceil kt \rceil}}{X_{(k)}} t^\xi \right), \quad 0 < t \leq 1.$$

By an abuse of notation let $\nu_n(y) := \nu_n(y^{-\xi}, \infty]$. Then for $0 < t \leq 1$,

$$\nu_n^{\leftarrow}(t) := \inf\{y : \nu_n(y^{-\xi}, \infty] \geq t\} = \inf \left\{ y : \sum_{i=1}^n \epsilon_{X_i/b(n/k)}(y^{-\xi}, \infty] \geq kt \right\} = \left(\frac{X_{\lceil kt \rceil}}{b(n/k)} \right)^{-1/\xi}.$$

Note that we can apply Vervaat's lemma (Proposition 3.3 in (Resnick, 2007, p.59)) to (3.8) to get

$$\sqrt{k} \left(\left(\frac{X_{\lceil kt \rceil}}{b(n/k)} \right)^{-1/\xi} - t \right) \Rightarrow W(t) \quad \text{in } D_l(0, 1]. \quad (3.9)$$

Therefore, using the continuous map $f : D_l(0, 1] \rightarrow D_l(0, 1]$ with $f(x)(t) = x(t)/t$, we have

$$\sqrt{k} \left(\left(\frac{X_{\lceil kt \rceil}}{b(n/k)} t^\xi \right)^{-1/\xi} - 1 \right) \Rightarrow \frac{W(t)}{t} \quad \text{in } D_l(0, 1]. \quad (3.10)$$

Also observe that

$$\begin{aligned} \sqrt{k} \log \left(\frac{X_{\lceil kt \rceil}}{b(n/k)} t^\xi \right) &= -\sqrt{k} \xi \log \left[1 - \left(1 - \left(\frac{X_{\lceil kt \rceil}}{b(n/k)} t^\xi \right)^{-1/\xi} \right) \right] \\ &= -\sqrt{k} \xi \left(\left(\frac{X_{\lceil kt \rceil}}{b(n/k)} t^\xi \right)^{-1/\xi} - 1 \right) + o_P \left(\sqrt{k} \xi \left(\left(\frac{X_{\lceil kt \rceil}}{b(n/k)} t^\xi \right)^{-1/\xi} - 1 \right) \right). \end{aligned} \quad (3.11)$$

So from (3.10) and (3.11) it follows that

$$\sqrt{k} \log \left(\frac{X_{\lceil kt \rceil}}{b(n/k)} t^\xi \right) \Rightarrow -\xi \frac{W(t)}{t} \quad \text{in } D_l(0, 1]. \quad (3.12)$$

We again use the continuous mapping theorem with $f : D_l(0, 1] \rightarrow D_l(0, 1]$ defined as $f(x)(t) = x(t) - x(1)$ to get the following:

$$\begin{aligned} \sqrt{k} \log \left(\frac{X_{\lceil kt \rceil}}{X_{(k)}} t^\xi \right) &= \sqrt{k} \log \frac{X_{\lceil kt \rceil}}{X_{(k)}} t^\xi - \sqrt{k} \log \frac{X_{(k)}}{b(n/k)} \\ &\Rightarrow \xi W(1) - \xi \frac{W(t)}{t} \quad \text{in } D_l(0, 1]. \end{aligned} \quad (3.13)$$

We know that $tW(1) - W(t) \stackrel{d}{=} B(t)$ on $D_l[0, 1]$. Therefore it is true on a restriction and hence

$$\sqrt{k} \left(\log \frac{X_{\lceil kt \rceil}}{X_{(k)}} + \xi \log t \right) \Rightarrow \xi t^{-1} B(t) \quad \text{in } D_l(0, 1]$$

where B is a Brownian Bridge on $[0, 1]$ restricted to $(0, 1]$. Furthermore, we also get

$$\left(-\log \frac{\lceil kt \rceil}{k}, -\xi \log \frac{\lceil kt \rceil}{k} + \sqrt{k} \left(\log \frac{X_{\lceil kt \rceil}}{X_{(k)}} + \xi \log \frac{\lceil kt \rceil}{k} \right) \right) \Rightarrow \left(-\log t, -\xi \log t + \xi \frac{B(t)}{t} \right) \quad \text{in } D_l^2(0, 1] \quad (3.14)$$

using the converging together lemma (Proposition 3.1 in (Resnick, 2007, p.57)) and the fact that

$$\sqrt{k} \left(\log \frac{[kt]}{k} - \log t \right) \rightarrow 0$$

locally uniformly on $(0, 1]$. The weak convergence of the set \mathcal{QN}_n follows easily from (3.14). We give the argument in the proof of Theorem 4.3 and do not give it here to avoid repetition. \square

4. Limit results for the ME Plots

4.1. Empirical ME function for known distribution F

Suppose X_1, \dots, X_n is an i.i.d. sample from distribution F . Yang (1978) studied the properties the empirical ME function $\hat{M}(u)$ in (1.2) as an estimator of $M(u)$. She showed that $\hat{M}(u)$ is uniformly strongly consistent for $M(u)$: for any $0 < b < \infty$

$$P \left[\lim_{n \rightarrow \infty} \sup_{0 \leq u \leq b} \left| \hat{M}(u) - M(u) \right| = 0 \right] = 1.$$

Yang (1978) also proved a weak limit for $\hat{M}(u)$: for any $0 < b < 1$

$$\sqrt{n} \left(\hat{M}(F^{\leftarrow}(t)) - M(F^{\leftarrow}(t)) \right) \Rightarrow U(t)$$

where $U(t)$ is a Gaussian process on $[0, b]$ with covariance function

$$\Gamma(s, t) = \frac{(1-t)\sigma^2(t) - t\theta^2(t)}{(1-s)(1-t)^2} \quad \text{for all } 0 \leq s \leq t \leq b$$

with

$$\sigma^2(t) = \text{var} \left(X I_{[t < F(X) \leq 1]} \right) \quad \text{and} \quad \theta(t) = E \left(X I_{[t < F(X) \leq 1]} \right).$$

Although these properties are stated for the empirical ME function it can be shown using the arguments at the end of the proof of Theorem 4.3 that the ME plots also exhibit the same features when the distribution F is known.

4.2. ME plot in the regularly varying case

Yang (1978) does not explain the behavior of $\hat{M}(u)$ near the right end point of F . Here we study the asymptotic properties of the ME plot when the explicit form of the distribution F is not known. As we had discussed for the QQ plots case, suppose we want to test for the null hypothesis that $\bar{F} \in RV_{1/\xi}$ for some $\xi > 0$. Ghosh and Resnick (2010) proved the limit in probability of a suitably scaled version of the ME plot under this null hypothesis:

Theorem 4.1. *If X_1, \dots, X_n are i.i.d. observations with distribution F satisfying $\bar{F} \in RV_{-1/\xi}$ with $0 < \xi < 1$, then in \mathcal{F}*

$$\mathcal{S}_n := \frac{1}{X_{(k)}} \left\{ (X_{(i)}, \hat{M}(X_{(i)})) : i = 2, \dots, k \right\} \xrightarrow{P} \mathcal{S} := \left\{ \left(t, \frac{\xi}{1-\xi} t \right) : t \geq 1 \right\}. \quad (4.1)$$

In this paper we obtain the weak limit of the ME plot when the null hypothesis that $\bar{F} \in RV_{1/\xi}$ for some $\xi > 0$ holds. The weak limit depends on the value of ξ . We get different limits depending on whether $\xi \leq 1/2$, $1/2 < \xi < 1$ or $\xi \geq 1$.

4.2.1. Case I: $0 < \xi < 1/2$

In this case $\text{var}(X_1)$ exists and we obtain a Gaussian limit for the suitably normalized ME plots. The following assumption is essential to prove the limit. This is comparable to the Assumption 3.2 that was required to obtain the weak limit of the QQ plot. As we discussed in Remark 3.5 it is very difficult to check this condition in practice.

Assumption 4.2. We say that F satisfies condition R' if

$$\lim_{n \rightarrow \infty} \sqrt{k} \left(\frac{n}{k} \bar{F}(b(n/k)y^{-\xi}) - y \right) = 0$$

locally uniformly on $(0, \infty]$ and

$$\sqrt{k} \int_1^\infty \left| \frac{n}{k} \bar{F}(b(n/k)y) - y^{-1/\xi} \right| dy \rightarrow 0$$

as $k, n, n/k \rightarrow \infty$.

Theorem 4.3. Suppose X_1, \dots, X_n are i.i.d. observations from distribution F satisfying $\bar{F} \in RV_{-1/\xi}$ with $0 < \xi < 1/2$ and condition R' . Then

$$\begin{aligned} \mathcal{MN}_n &:= \left\{ \left(\left(\frac{i}{k} \right)^{-\xi}, \frac{\xi}{1-\xi} \left(\frac{i}{k} \right)^{-\xi} \right) \right. \\ &\quad \left. + \sqrt{k} \left(\frac{X_{(i)}}{X_{(k)}} - \left(\frac{i}{k} \right)^{-\xi}, \frac{\hat{M}(X_{(i)})}{X_{(k)}} - \frac{\xi}{1-\xi} \left(\frac{i}{k} \right)^{-\xi} \right) : i = 2, \dots, k \right\} \\ &\Rightarrow \mathcal{MN} := \left\{ \left(t^{-\xi} + \xi t^{-(1+\xi)} B(t), \frac{\xi}{1-\xi} t^{-\xi} + \xi t^{-1} \int_0^t y^{-(1+\xi)} B(y) dy \right), 0 < t \leq 1 \right\} \quad \text{in } \mathcal{F}, \end{aligned}$$

where $B(t)$ is the standard Brownian bridge on $[0, 1]$ restricted to $(0, 1]$.

Remark 4.4. Similar to Theorem 3.3 we look at the ME plot as the probability limit perturbed by the normalized deviation around it and obtain a weak limit in Theorem 4.3. The assumption that $\xi < 1/2$ is essential. Note that

$$\int_0^t y^{-(1+\xi)} W(y) dy = \int_{t^{-\xi}}^\infty W(u^{-1/\xi}) du = \int_{t^{-\xi}}^\infty \int_0^{y^{-1/\xi}} dW(s) dy = \int_0^t s^{-\xi} dW(s).$$

and it well known that the last integral exists if and only if $\int_0^t s^{-2\xi} ds < \infty$ for which it is necessary and sufficient to have $\xi < 1/2$, cf. (Øksendal, 2003, Lemma 3.1.5, p.26). This means

$$\int_0^t y^{-(1+\xi)} B(y) dy \stackrel{d}{=} \int_0^t y^{-(1+\xi)} W(y) dy - W(1) \int_0^t y^{-\xi} dy$$

exists if and only if $\xi < 1/2$ and the same is true for the limit \mathcal{MN} .

Proof. Consider a functional form of the ME plot:

$$S_n = \left(S_n^{(1)}, S_n^{(2)} \right) := \left(\frac{X_{(\lceil kt \rceil)}}{X_{(k)}}, \frac{\hat{M}(X_{(\lceil kt \rceil)})}{X_{(k)}} \right), \quad t \in (0, 1] \quad (4.2)$$

as random elements in $D_l^2(0, 1]$. Following the proof of Theorem 3.2 in Ghosh and Resnick (2010) we know that $S_n(\cdot) \xrightarrow{P} S(\cdot)$ in $D_l^2(0, 1]$, where

$$S(t) = \left(t^{-\xi}, \frac{\xi}{1-\xi} t^{-\xi} \right) \quad \text{for all } 0 < t \leq 1.$$

Applying Vervaat's lemma (Resnick, 2007, Proposition 3.3, p.59) to (3.8) we get

$$\left(\sqrt{k} \left(\left(\frac{X_{(\lceil kt \rceil)}}{b(n/k)} \right)^{-1/\xi} - t \right), \sqrt{k}(\nu_n(y^{-\xi}, \infty] - y) \right) \Rightarrow (-W(t), W(t)) \quad \text{in } D_l^2(0, 1]. \quad (4.3)$$

By an application of the Delta-method to (4.3) we obtain a weak limit of the first component

$$\sqrt{k} \left(\frac{X_{(\lceil kt \rceil)}}{b(n/k)} - t^{-\xi} \right) \Rightarrow \xi t^{-(1+\xi)} W(t) \quad \text{in } D_l(0, 1]. \quad (4.4)$$

This means that

$$\sqrt{k} (S_n^{(1)}(t) - S^{(1)}(t)) \quad (4.5)$$

$$\begin{aligned} &= \sqrt{k} \left(\frac{X_{(\lceil kt \rceil)}}{X_{(k)}} - t^{-\xi} \right) = \frac{b(n/k)}{X_{(k)}} \sqrt{k} \left(\frac{X_{(\lceil kt \rceil)}}{b(n/k)} - t^{-\xi} \right) - \frac{b(n/k)}{X_{(k)}} \sqrt{k} \left(\frac{X_{(k)}}{b(n/k)} - 1 \right) t^{-\xi} \\ &\Rightarrow \xi t^{-(1+\xi)} W(t) - \xi t^{-\xi} W(1) \stackrel{d}{=} -\xi t^{-(1+\xi)} B(t) \end{aligned} \quad (4.6)$$

Next we consider the second component of S_n and observe that

$$S_n^{(2)}(t) := \frac{\hat{M}(X_{(\lceil kt \rceil)})}{X_{(k)}} = \frac{k}{\lceil kt \rceil - 1} \int_{X_{(\lceil kt \rceil)}/b(n/k)}^{\infty} \hat{\nu}_n(x, \infty] dx,$$

where

$$\hat{\nu}_n(\cdot) := \frac{1}{k} \sum_{i=1}^n \epsilon_{X_i/X_{(k)}}(\cdot) \quad (4.7)$$

Using (3.8) and the converging together lemma (Resnick, 2007, Proposition 3.1, p.57) we know that

$$\left(\sqrt{k}(\nu_n(y, \infty] - y^{-1/\xi}), \frac{X_{(k)}}{b(n/k)} \right) \Rightarrow (W(y^{-1/\xi}), 1) \quad \text{in } D(0, \infty] \times \mathbb{R}.$$

Using the fact that the map $\hat{T} : D(0, \infty] \times \mathbb{R} \rightarrow D(0, \infty]$ defined by $\hat{T}(g, x)(t) = g(tx)$ is continuous at $g \in C(0, \infty]$ we get

$$\begin{aligned} &\sqrt{k} \left(\hat{\nu}_n(y, \infty] - \left(y \frac{X_{(k)}}{b(n/k)} \right)^{-1/\xi} \right) \\ &= \sqrt{k} \left(\nu_n \left(y \frac{X_{(k)}}{b(n/k)}, \infty \right] - \left(y \frac{X_{(k)}}{b(n/k)} \right)^{-1/\xi} \right) \Rightarrow W(y^{-1/\xi}) \quad \text{in } D(0, \infty]. \end{aligned}$$

Furthermore, (3.8) implies

$$H_n(y) := \sqrt{k} \left(\hat{\nu}_n(y, \infty] - y^{-1/\xi} \right) \Rightarrow W(y^{-1/\xi}) - y^{-1/\xi} W(1) \stackrel{d}{=} B(y^{-1/\xi}) =: H(y). \quad (4.8)$$

Define the maps T and T_K from $D[0, \infty)$ to $D[1, \infty)$ by

$$T(f)(t) = \int_t^\infty f(x) dx \quad \text{and} \quad T_K(f)(t) = \int_t^{K \vee t} f(x) dx. \quad (4.9)$$

We understand $T(f)(t) = \infty$ if f is not integrable. By (4.8) and the continuity of the map T_K we get that $T_K(H_n) \Rightarrow T_K(H)$. We also claim that for any $\epsilon > 0$

$$\lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P \left[\|T_K(H_n) - T(H_n)\| > \epsilon \right] = 0. \quad (4.10)$$

Note that for any $\epsilon > 0$

$$\begin{aligned}
& \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P \left[\|T_K(H_n) - T(H_n)\| > \epsilon \right] \\
& \leq \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P \left[\sqrt{k} \left| \int_K^\infty (\hat{\nu}_n(y, \infty) - y^{-1/\xi}) dy \right| > \epsilon \right] \\
& \leq \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P \left[\sqrt{k} \left| \int_K^\infty \left(\nu_n \left(y \frac{X_{(k)}}{b(n/k)}, \infty \right) - \left(y \frac{X_{(k)}}{b(n/k)} \right)^{-1/\xi} \right) dy \right| > \epsilon/2 \right] \\
& \quad + \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P \left[\sqrt{k} \left| \int_K^\infty y^{-1/\xi} \left(\left(\frac{X_{(k)}}{b(n/k)} \right)^{-1/\xi} - 1 \right) dy \right| > \epsilon/2 \right]
\end{aligned}$$

Using (3.9) and the assumption that $\xi < 1/2$ we get

$$\lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P \left[\sqrt{k} \left| \int_K^\infty y^{-1/\xi} \left(\left(\frac{X_{(k)}}{b(n/k)} \right)^{-1/\xi} - 1 \right) dy \right| > \epsilon/2 \right] = 0.$$

Using a change of variable we obtain

$$\begin{aligned}
& \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P \left[\sqrt{k} \left| \int_K^\infty \left(\nu_n \left(y \frac{X_{(k)}}{b(n/k)}, \infty \right) - \left(y \frac{X_{(k)}}{b(n/k)} \right)^{-1/\xi} \right) dy \right| > \epsilon/2 \right] \\
& \leq \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P \left[\sqrt{k} \left| \int_{K X_{(k)}/b(n/k)}^\infty \left(\nu_n(u, \infty) - u^{-1/\xi} \right) \frac{b(n/k)}{X_{(k)}} du \right| > \epsilon/2 \right]
\end{aligned}$$

Now fix any $\eta > 0$ and note that

$$\lim_{n \rightarrow \infty} P \left[\left| \frac{X_{(k)}}{b(n/k)} - 1 \right| > \eta \text{ or } \left| \frac{b(n/k)}{X_{(k)}} - 1 \right| > \eta \right] = 0.$$

Therefore,

$$\begin{aligned}
& \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P \left[\sqrt{k} \left| \int_{K X_{(k)}/b(n/k)}^\infty \left(\nu_n(u, \infty) - u^{-1/\xi} \right) \frac{b(n/k)}{X_{(k)}} du \right| > \epsilon/2 \right] \\
& \leq \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P \left[(1 + \eta) \sqrt{k} \left| \int_{K(1-\eta)}^\infty \left(\nu_n(u, \infty) - u^{-1/\xi} \right) du \right| > \epsilon/2 \right] + o(1).
\end{aligned}$$

Now since F satisfies condition R' it suffices to show that

$$\lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P \left[\sqrt{k} \left| \int_{K(1-\eta)}^\infty \left(\nu_n(y, \infty) - \frac{n}{k} \bar{F}(b(n/k)y) \right) dy \right| > \frac{\epsilon}{2(1-\eta)} \right] = 0 \quad (4.11)$$

This can be easily proved using the arguments in the proof of Proposition 9.1 in (Resnick, 2007, p.296). Observe that

$$\begin{aligned}
& P \left[\sqrt{k} \left| \int_{K(1-\eta)}^\infty \left(\nu_n(y, \infty) - \frac{n}{k} \bar{F}(b(n/k)y) \right) dy \right| > \frac{\epsilon}{2(1-\eta)} \right] \\
& \leq P \left[\frac{1}{\sqrt{k}} \int_{K(1-\eta)}^\infty \sum_{i=1}^n |\epsilon_{X_i/b(n/k)}(y, \infty) - \bar{F}(b(n/k)y)| dy > \frac{\epsilon}{2(1-\eta)} \right]
\end{aligned}$$

$$\begin{aligned}
&\leq \left(\frac{\epsilon}{2(1-\eta)}\right)^{-2} \int_{K(1-\eta)}^{\infty} \frac{n}{k} \text{var}[\epsilon_{X_i/b(n/k)}(y, \infty)] dy \\
&\leq \left(\frac{\epsilon}{2(1-\eta)}\right)^{-2} \int_{K(1-\eta)}^{\infty} \frac{n}{k} \bar{F}(b(n/k)y) dy \xrightarrow{n \rightarrow \infty} \left(\frac{\epsilon}{2(1-\eta)}\right)^{-2} \int_{K(1-\eta)}^{\infty} y^{-1/\xi} dy.
\end{aligned}$$

The last limit follows from Karamata's Theorem, c.f. (Resnick, 2007, p.25). Since $\xi < 1$, the integral in the last expression is finite and therefore (4.11), and hence (4.10) holds. From Theorem 3.5 in (Resnick, 2007, p.56) we get $T(H_n) \Rightarrow T(H) = \int_t^\infty B(y^{-1/\xi}) dy$ in $C(0, \infty]$.

Now consider the random element Y_n in the space $D_l(0, 1] \times C(0, \infty]$,

$$Y_n := \left(\frac{X_{(\lceil k \cdot \rceil)}}{X_{(k)}}, T(H_n) \right).$$

From what we have obtained so far it is easy to check that $Y_n \Rightarrow Y$, where

$$Y(t, u) = \left(t^{-\xi}, \int_u^\infty B(y^{-1/\xi}) dy \right).$$

The map $\tilde{T} : D_l(0, 1] \times C(0, \infty] \rightarrow D_l(0, 1]$ defined by

$$\tilde{T}(f, g)(t) = g(f(t)) \quad \text{for all } 0 < t \leq 1$$

is continuous at $(f, g) \in C(0, 1] \times C(0, \infty]$ and therefore

$$\tilde{T}(Y_n)(t) = \sqrt{k} \int_{X_{(\lceil kt \rceil)}/X_{(k)}}^\infty \left(\hat{\nu}_n(y, \infty) - y^{-1/\xi} \right) dy \Rightarrow \int_{t^{-\xi}}^\infty B(y^{-1/\xi}) dy \quad \text{in } D_l(0, 1].$$

This gives us the weak limit for the second component of S_n

$$\sqrt{k}(S_n^{(2)}(t) - S^{(2)}(t)) = \sqrt{k} \left(\frac{\hat{M}(X_{(\lceil kt \rceil)})}{X_{(k)}} - \frac{\xi}{1-\xi} t^{-\xi} \right) \Rightarrow \frac{1}{t} \int_{t^{-\xi}}^\infty B(y^{-1/\xi}) dy \quad \text{in } D_l(0, 1]. \quad (4.12)$$

We proved the weak convergence of the two components of S_n separately. Both these limits were obtained using transformations on ν_n and the weak limit of ν_n . Hence we get the weak convergence of S_n

$$\sqrt{k}(S_n(t) - S(t)) \Rightarrow \left(\xi t^{-(1+\xi)} B(t), \frac{1}{t} \int_{t^{-\xi}}^\infty B(y^{-1/\xi}) dy \right) \quad \text{in } D_l^2(0, 1].$$

It is then easy to check that

$$\left(\xi t^{-(1+\xi)} B(t), \frac{1}{t} \int_{t^{-\xi}}^\infty B(y^{-1/\xi}) dy \right) \stackrel{d}{=} \left(\xi t^{-(1+\xi)} B(t), \frac{\xi}{t} \int_0^t y^{-(1+\xi)} B(y) dy \right) \quad \text{in } D_l^2(0, 1].$$

Also observe that

$$\begin{aligned}
\tilde{S}_n(t) &:= \left(\left(\frac{\lceil kt \rceil}{k} \right)^{-\xi}, \frac{\xi}{1-\xi} \left(\frac{\lceil kt \rceil}{k} \right)^{-\xi} \right) \\
&\quad + \sqrt{k} \left(\frac{X_{(\lceil kt \rceil)}}{X_{(k)}} - \left(\frac{\lceil kt \rceil}{k} \right)^{-\xi}, \frac{\hat{M}(X_{(\lceil kt \rceil)})}{X_{(k)}} - \frac{\xi}{1-\xi} \left(\frac{\lceil kt \rceil}{k} \right)^{-\xi} \right) \\
&\Rightarrow \tilde{S}(t) := \left(t^{-\xi}, \frac{\xi}{1-\xi} t^{-\xi} \right) + \left(\xi t^{-(1+\xi)} B(t), \frac{\xi}{t} \int_0^t y^{-(1+\xi)} B(y) dy \right) \quad \text{in } D_l^2(0, 1],
\end{aligned}$$

since

$$\sqrt{k} \left(\left(\frac{[kt]}{k} \right)^{-\xi} - t^{-\xi} \right) \rightarrow 0 \quad \text{as } k \rightarrow \infty$$

locally uniformly on $(0, 1]$

Now we adapt this result to convergence of the ME plot and prove the statement of the theorem. By Lemma 2.3 it suffices to show that

$$\lim_{n \rightarrow \infty} E[h^\vee(\mathcal{MN}_n)] = E[h^\vee(\mathcal{MN})].$$

for any continuous function $h : \mathbb{R}^2 \mapsto \mathbb{R}_+$ with a compact support. So take any such function h . By the Skorokhod representation theorem there exists a probability space (Ω, \mathcal{G}, P) and random elements S_n^* and S^* in $D_l^2(0, 1]$ such that

$$S_n^* \stackrel{d}{=} \tilde{S}_n \quad \text{and} \quad S^* \stackrel{d}{=} \tilde{S}$$

in the sense of fidi and

$$S_n^* \rightarrow S^* \quad \text{in } D_l^2(0, 1] \quad P - a.s.$$

Now observe that

$$h^\vee(\mathcal{MN}) = \sup_{x \in \mathcal{MN}} h(x) \stackrel{d}{=} \sup_{0 < t \leq 1} h(S^*(t)) \quad \text{and} \quad h^\vee(\mathcal{MN}_n) = \sup_{x \in \mathcal{MN}_n} h(x) \stackrel{d}{=} \sup_{1/k < t \leq 1} h(S_n^*(t)).$$

Since $S^*(t)$ is continuous, $k \rightarrow \infty$ and h is continuous with a compact support, we get

$$\sup_{1/k < t \leq 1} h(S_n^*(t)) \rightarrow \sup_{0 < t \leq 1} h(S^*(t)) \quad P - a.s.$$

and since h is bounded we apply the dominated convergence theorem to get

$$E[h^\vee(\mathcal{MN}_n)] = E\left[\sup_{1/k < t \leq 1} h(S_n^*(t))\right] \rightarrow E\left[\sup_{0 < t \leq 1} h(S^*(t))\right] = E[h^\vee(\mathcal{MN})].$$

Hence the proof is complete. \square

4.3. Case II: $1/2 < \xi < 1$

When $1/2 < \xi < 1$, the distribution F admits a finite mean but not a finite variance. The ME function though exists and we know the limit in probability of the scaled ME plot from Theorem 4.1.

Assumption 4.5. We say that F satisfies condition R'' if

$$\lim_{n \rightarrow \infty} \sqrt{k} \left(\frac{n}{k} \bar{F}(b(n/k)y^{-\xi}) - y \right) = 0$$

locally uniformly on $(0, \infty]$ and

$$\frac{1}{b(n)} \left(\frac{kb(n/k)}{1 - \xi} u^{1-\xi} - C_{ku,n} \right) \rightarrow 0$$

for every $0 < u < 1$ where for any $l < n$

$$C_{l,n} := n \int_0^{l/n} F^{\leftarrow}(1 - u) du. \quad (4.13)$$

Theorem 4.6. Suppose X_1, \dots, X_n are i.i.d. observations from distribution F satisfying $\bar{F} \in RV_{-1/\xi}$ with $1/2 < \xi < 1$ and condition R'' . Then

$$\begin{aligned} \mathcal{MN}_n &:= \left\{ \left(\left(\frac{i}{k} \right)^{-\xi}, \frac{\xi}{1-\xi} \left(\frac{i}{k} \right)^{-\xi} \right) \right. \\ &\quad \left. + \left(\sqrt{k} \left(\frac{X_{(i)}}{X_{(k)}} - \left(\frac{i}{k} \right)^{-\xi} \right), \frac{kb(n/k)}{b(n)} \left(\frac{\hat{M}(X_{(i)})}{X_{(k)}} - \frac{\xi}{1-\xi} \left(\frac{i}{k} \right)^{-\xi} \right) \right) : i = 2, \dots, k \right\} \\ &\Rightarrow \mathcal{MN} := \left\{ \left(t^{-\xi} + \xi t^{-(1+\xi)} B(t), \frac{\xi}{1-\xi} t^{-\xi} + t^{-1} S_{1/\xi} \right), 0 < t \leq 1 \right\} \quad \text{in } \mathcal{F}, \end{aligned}$$

where $B(t)$ is the standard Brownian bridge on $[0, 1]$ restricted to $(0, 1]$ and $S_{1/\xi}$ is a stable random variable independent of $B(t)$ with characteristic function

$$E[e^{itS_{1/\xi}}] = \exp \left\{ -\frac{1}{1-\xi} \Gamma\left(2 - \frac{1}{\xi}\right) \cos \frac{\pi}{2\xi} |t|^{1/\xi} \left[1 - i \operatorname{sgn}(t) \tan \frac{\pi}{2\xi} \right] \right\} \quad (4.14)$$

Remark 4.7. A very interesting point to note here is that the two coordinates of the weak limit \mathcal{MN} are independent. The empirical ME function depends on the sum of the order statistics $X_{(1)}, \dots, X_{(k)}$. When $1/2 < \xi < 1$ this sum is dominated by a very few high order statistics and it turns out that the contribution of $X_{(k)}$ to the suitably normalized $\hat{M}(X_{(k)})$ vanishes in the limit. The proof below formalizes this idea.

This feature is in stark contrast to what happens in the case $0 < \xi < 1/2$. In that case all the top k order statistics have some contribution to $\hat{M}(X_{(k)})$ in the limit. Hence the two coordinates in the limit are obtained from the same Gaussian process and are definitely not independent.

Remark 4.8. Unfortunately, we are unable to obtain a proper weak limit of the ME plot in the case when $\xi = 1/2$. In this case it is known that the weak limit of the suitably normalized sum of the first k order statistics is Gaussian, cf. Csörgö et al. (1991). So this would be similar to what happens when $0 < \xi < 1/2$ but the problem is that the integral $\int_0^t y^{-2} dB(y)$ does not exist. It is possible to redefine the ME plot in a different way by leaving out a few of the top order statistics and obtain a limit in that case but we did not pursue that direction.

Proof. From Theorem 3 in Csörgö et al. (1986) we know that if $l = l_n \rightarrow \infty$ with $l_n/n \rightarrow 0$ then

$$\frac{1}{b(n)} \left(\sum_{i=1}^l X_{(i)} - C_{l,n} \right) \Rightarrow S_{1/\xi}. \quad (4.15)$$

Observe that by Karamata's theorem (Resnick, 2007, Theorem 2.1, p.25)

$$C_{l,n} = n \int_{n/l}^{\infty} b(s)/s^2 ds \sim n \frac{(n/l)b(n/l)}{(n/l)^2(1-\xi)} = \frac{lb(n/l)}{1-\xi}$$

Choose $l = l_n$ such that $l/k \rightarrow 0$ as $n \rightarrow \infty$. Fix any $0 < u < 1$. Then

$$\begin{aligned} V_n(t) &:= (V_n^{(1)}(t), V_n^{(2)}(t)) = \left(\sqrt{k} \left(\frac{X_{(\lceil kt \rceil)}}{X_{(k)}} - t^{-\xi} \right), \frac{1}{b(n)} \left(\sum_{i=1}^l X_{(i)} - C_{l,n} \right) \right) \\ &\Rightarrow (\xi t^{-(1+\xi)} B(t), S_{1/\xi}) \quad \text{in } D_l[u, 1] \times \mathbb{R} \end{aligned} \quad (4.16)$$

where $B(t)$ and $S_{1/\xi}$ are as described in the statement of the theorem. The convergence of the coordinates $V_n^{(1)}(t)$ and $V_n^{(2)}$ of $V_n(t)$ follows from (4.5) and (4.15). The asymptotic independence of $V_n^{(1)}(t)$ and $V_n^{(2)}$ is

a consequence of Theorem D in [Csörgo and Mason \(1985\)](#) or Satz 4 in [Rossberg \(1967\)](#). Using (3.8) we get that

$$\sqrt{k} \left(\frac{1}{kb(n/k)} \sum_{i=\lceil ku \rceil+1}^{\lceil kt \rceil} X_{(i)} - \frac{1}{1-\xi} (t^{1-\xi} - u^{1-\xi}) \right) \Rightarrow \int_u^t W(y) dy \quad \text{in } D_l[u, 1]$$

and since $\xi > 1/2$, $kb(n/k)/(b(n)\sqrt{k}) \rightarrow 0$ which implies

$$U_n^{(2)}(t) := \frac{kb(n/k)}{b(n)} \left(\frac{1}{kb(n/k)} \sum_{i=\lceil ku \rceil+1}^{\lceil kt \rceil} X_{(i)} - \frac{1}{1-\xi} (t^{1-\xi} - u^{1-\xi}) \right) \rightarrow \mathbf{0} \quad \text{in } D_l[u, 1] \quad (4.17)$$

where $\mathbf{0} \in D_l[u, 1]$ denotes the identically zero function. Furthermore, using Theorem 2 in [Csörgo et al. \(1986\)](#) we get

$$\frac{1}{\sqrt{k}b(n/k)} \left(\sum_{i=l+1}^{\lceil ku \rceil} X_{(i)} - (C_{ku,n} - C_{l,n}) \right) \Rightarrow N(0, 1)$$

and hence

$$U_n^{(3)} := \frac{1}{b(n)} \left(\sum_{i=l+1}^{\lceil ku \rceil} X_{(i)} - (C_{ku,n} - C_{l,n}) \right) \rightarrow 0. \quad (4.18)$$

Combining (4.16), (4.17) and (4.18) and the converging together lemma ([Resnick, 2007](#), Proposition 3.1, p.57), we get an important building block of this proof

$$U_n(t) := (V_n^{(1)}(t), U_n^{(2)}(t), U_n^{(3)}, V_n^{(2)}) \Rightarrow (\xi t^{-(1+\xi)} W(t), \mathbf{0}, 0, S_{1/\xi}) \quad \text{in } D_l^2[u, 1] \times \mathbb{R}^2. \quad (4.19)$$

Next we consider

$$Z_n(t) := \left(\sqrt{k} \left(\frac{X_{(\lceil kt \rceil)}}{X_{(k)}} - t^{-\xi} \right), \frac{kb(n/k)}{b(n)} \left(\frac{\hat{M}(X_{(\lceil kt \rceil)})}{X_{(k)}} - \frac{\xi}{1-\xi} t^{-\xi} \right) \right) \in D_l^2[u, 1]$$

and focus on the second coordinate $Z_n^{(2)}(t)$:

$$\begin{aligned} Z_n^{(2)}(t) &= \frac{kb(n/k)}{b(n)} \left(\frac{\hat{M}(X_{(\lceil kt \rceil)})}{X_{(k)}} - \frac{\xi}{1-\xi} t^{-\xi} \right) \\ &= \frac{kb(n/k)}{b(n)} \left(\frac{\hat{M}(X_{(\lceil kt \rceil)})}{b(n/k)} - \frac{\xi}{1-\xi} t^{-\xi} \right) + o_P(1) \\ &= \frac{kb(n/k)}{b(n)} \left(\frac{1}{(\lceil kt \rceil - 1)b(n/k)} \sum_{i=1}^{\lceil kt \rceil - 1} X_{(i)} - \frac{X_{(\lceil kt \rceil)}}{b(n/k)} - \frac{\xi}{1-\xi} t^{-\xi} \right) + o_P(1) \\ &= \frac{kb(n/k)}{b(n)} \left(\frac{1}{(\lceil kt \rceil - 1)b(n/k)} \sum_{i=1}^{\lceil kt \rceil - 1} X_{(i)} - \frac{1}{1-\xi} t^{-\xi} \right) + o_P(1) \\ &= \frac{kb(n/k)}{tb(n)} \left(\frac{1}{kb(n/k)} \sum_{i=1}^{\lceil kt \rceil - 1} X_{(i)} - \frac{1}{1-\xi} t^{1-\xi} \right) + o_P(1) \\ &= \frac{1}{t} U_n^{(2)}(t) + \frac{1}{t} U_n^{(3)} + \frac{1}{t} V_n^{(2)} + o_P(1) \end{aligned}$$

where the last equality holds because of condition R'' . Therefore, we get

$$Z_n(t) \Rightarrow (\xi t^{-(1+\xi)} B(t), t^{-1} S_{1/\xi}) \quad \text{in } D_l^2[u, 1].$$

Since the above limit holds for every $0 < u < 1$ it holds in $D_l(0, 1]$ as well. The rest of the proof is completed following the last part of the proof of Theorem 4.3. \square

4.4. Case III: $\xi \geq 1$

In this case the distribution F need not have a finite mean and the ME function may not be defined. It definitely does not exist if $\xi > 1$. Still the empirical ME plot can have a limit.

Theorem 4.9. Assume X_1, \dots, X_n are i.i.d. observations with distribution F such that $\bar{F} \in RV_{-1/\xi}$ and condition R' holds:

1. If $\xi > 1$, then

$$\begin{aligned} \mathcal{MN}_n &:= \left\{ \left(\frac{i}{k} \right)^{-\xi} + \sqrt{k} \left(\frac{X_{(i)}}{X_{(k)}} - \left(\frac{i}{k} \right)^{-\xi} \right), \frac{\hat{M}(X_{(i)})}{b(n)/k} : i = 2, \dots, k \right\} \\ &\Rightarrow \mathcal{MN} := \left\{ \left(\xi t^{-(1+\xi)} B(t), t S_{1/\xi} \right) : t \geq 1 \right\} \end{aligned}$$

in \mathcal{F} , where $S_{1/\xi}$ is the positive stable random variable with index $1/\xi$ which satisfies for $t \in \mathbb{R}$

$$E[e^{itS_{1/\xi}}] = \exp \left\{ -\Gamma \left(1 - \frac{1}{\xi} \right) \cos \frac{\pi}{2\xi} |t|^{1/\xi} \left[1 - i \operatorname{sgn}(t) \tan \frac{\pi}{2\xi} \right] \right\}$$

and $B(t)$ is a Brownian bridge independent of $S_{1/\xi}$.

2. If $\xi = 1$ and k satisfies $k = k(n) \rightarrow \infty$, $k/n \rightarrow 0$, and

$$kb(n/k)/b(n) \rightarrow 1 \quad \text{as } n \rightarrow \infty,$$

then

$$\begin{aligned} \mathcal{MN}_n &:= \left\{ \left(\frac{i}{k} \right)^{-\xi} + \sqrt{k} \left(\frac{X_{(i)}}{X_{(k)}} - \left(\frac{i}{k} \right)^{-\xi} \right), \frac{\hat{M}(X_{(i)})}{b(n/k)} - \frac{kC_{k,n}^*}{ib(n)} : i = 2, \dots, k \right\} \\ &\Rightarrow \mathcal{MN} := \left\{ t \left(t^{-1} B(t), S_1 - 1 - \log t \right) : t \geq 1 \right\} \end{aligned}$$

in \mathcal{F} , where

$$C_{k,n}^* = n \int_{1/n}^{k/n} F^{\leftarrow}(1-u) du,$$

S_1 is a positively skewed stable random variable satisfying

$$E[e^{itS_1}] = \exp \left\{ it \int_0^\infty \left(\frac{\sin x}{x^2} - \frac{1}{x(1+x)} \right) dx - |t| \left[\frac{\pi}{2} + i \operatorname{sgn}(t) \log |t| \right] \right\}.$$

and $B(t)$ is a Brownian bridge independent of S_1 .

Proof. The theorem is proved in the same fashion as the previous ones. First we prove the weak limit in the functional form of the ME plot and then we infer the the weak limit of the plot as a random set. Define

$$S_n(t) = \begin{cases} \left(\sqrt{k} \left(\frac{X_{(\lceil kt \rceil)} - t^{-\xi}}{X_{(k)}} \right), \frac{\hat{M}(X_{(\lceil kt \rceil)})}{b(n)/k} \right) & \text{in part (i)} \\ \left(\sqrt{k} \left(\frac{X_{(\lceil kt \rceil)} - t^{-\xi}}{X_{(k)}} \right), \frac{\hat{M}(X_{(\lceil kt \rceil)})}{b(n)/k} - \frac{kC_{k,n}^*}{\lceil kt \rceil b(n)} \right) & \text{in part (ii)} \end{cases} \quad \text{for all } 0 < t \leq 1.$$

We have already proved the weak limit of $S_n^{(1)}(t)$ and the weak limit of $S_n^{(2)}(t)$ is proved in Theorem 3.4 in [Ghosh and Resnick \(2010\)](#). The rest of the proof is completed following the lines of Theorem 4.6. \square

5. Confidence Bounds for the Plots

5.1. QQ plots

The limit distribution we obtained in Theorem 3.3 is a linear transformation of $\{B(t)/t : 0 < t \leq 1\}$ where $B(\cdot)$ is a Brownian Bridge on $[0, 1]$. Clearly $B(t)/t$ blows up near $t = 0$. Thus for proper inference purposes in terms of constructing confidence bounds we need to truncate away from zero. Hence we will fix $0 < \epsilon < 1$ but very close to zero. Clearly again $\sqrt{k}(G_n(t) - \frac{1}{\alpha}G(t))$ converges weakly to $B(t)/t$ in $D_l[\epsilon, 1]$, following notations in Section 3. So in the limit we want to find $M > 0$ such that the probability of $P\left(\max_{\epsilon \leq t \leq 1} \frac{|B(t)|}{t} > M\right)$ is very small.

Note that it is easy to check that $W(t) := (1+t)B\left(\frac{t}{t+1}\right)$, $0 \leq t < \infty$ is a Brownian motion on $[0, \infty)$ when $B(t)$ is a Brownian Bridge on $[0, 1]$. Thus

$$\begin{aligned} \max_{\epsilon \leq t \leq 1} \frac{|B(t)|}{t} &= \max_{\epsilon \leq \frac{t}{t+1} \leq 1} \frac{|B(\frac{t}{t+1})|}{\frac{t}{t+1}} = \max_{\epsilon \leq \frac{t}{t+1} \leq 1} \frac{|W(t)|}{t} \\ &= \max_{\frac{\epsilon}{1-\epsilon} \leq t < \infty} \frac{|W(t)|}{t} = \max_{t \geq \delta} \frac{|W(t)|}{t}, \end{aligned}$$

where $\delta = \frac{\epsilon}{1-\epsilon}$.

Now we compute the above probability which helps us to calculate the probability that a straight line through the origin (which is the intended limit of the QQ plot) actually lies within the confidence bounds that we can build with the distributional result we obtained in Theorem 3.3.

Proposition 5.1. *Suppose $W(t)$ is a Standard Brownian Motion on $[0, \infty)$. Then for all $\delta > 0$ and $M > 0$, the following holds:*

$$P\left(\sup_{t \geq \delta} \frac{|W(t)|}{t} > M\right) = 4 \sum_{k=1}^{\infty} \left[\Phi((4k+1)M\sqrt{\delta}) - \Phi((4k-1)M\sqrt{\delta}) \right].$$

Proof. The proof is an application of a result in [Doob \(1949\)](#), see ([Shorack and Wellner, 1986](#), page 38). It says that if $W(\cdot)$ is a simple Brownian motion on $[0, \infty)$, then for $a, c \geq 0$ and $\alpha, \beta \geq 0$,

$$P(-(\alpha t + \beta) \leq W(t) \leq at + b, \forall t \geq 0) = 1 - \sum_{k=1}^{\infty} [e^{-2A_k} + e^{-2B_k} - e^{-2C_k} - e^{-2D_k}] \quad (5.1)$$

where

$$A_k = k^2 ab + (k-1)^2 \alpha \beta + k(k-1)(a\beta + b\alpha), \quad B_k = (k-1)^2 ab + k^2 \alpha \beta + k(k-1)(a\beta + b\alpha),$$

$$C_k = k^2(ab + \alpha\beta) + k(k-1)a\beta + k(k+1)b\alpha, \quad D_k = k^2(ab + \alpha\beta) + k(k+1)a\beta + k(k-1)b\alpha.$$

Now

$$\begin{aligned} & P\left(\sup_{t \geq \delta} \frac{|W(t)|}{t} > M\right) \\ &= 1 - P(-Mt \leq W(t) \leq M(t), \quad \forall t \geq \delta) \\ &= 1 - P(-M(t-\delta) - M\delta - W(\delta) \leq W(t) - W(\delta) \leq M(t-\delta) + M\delta - W(\delta), \quad \forall t \geq \delta) \\ &= 1 - \int_{s=-M\delta}^{M\delta} P(-Mt - (M\delta + s) \leq W(t) - W(\delta) \leq Mt + (M\delta - s) \quad \forall t \geq 0 | W(\delta) = s) f_Z(s) ds \end{aligned} \quad (5.2)$$

(conditioning on $W(\delta)$ and denoting the p.d.f. of $Z := W(\delta) \sim N(0, \delta)$ by f_Z)

$$= 1 - \int_{s=-M\delta}^{M\delta} P(-Mt - (M\delta + s) \leq W(t) \leq Mt + (M\delta - s) \quad \forall t \geq 0) f_Z(s) ds. \quad (5.3)$$

The last equality results from the stationarity and independent increment property of B.M., $\{W(t) - W(\delta) : t \geq \delta\} \perp W(\delta)$ and $\{W(t) - W(\delta) : t \geq \delta\} \stackrel{d}{=} \{W(t) : t \geq 0\}$. Now we can use (5.1) on the integrand of (5.3) and calculate

$$P\left(\sup_{t \geq \delta} \frac{|W(t)|}{t} > M\right) = 1 - \int_{s=-M\delta}^{M\delta} \left[1 - \sum_{k=1}^{\infty} (e^{-2A_k} + e^{-2B_k} - e^{-2C_k} - e^{-2D_k})\right] f_Z(s) ds.$$

where

$$\begin{aligned} A_k &= E_k^2 \delta - E_k s, & B_k &= E_k^2 \delta + E_k s, \\ C_k &= F_k^2 \delta - F_k s, & D_k &= F_k^2 + F_k s, \end{aligned}$$

with $E_k = (2k-1)M$, $F_k = 2kM$, $\forall k \geq 1$. Also note that for any $a, b \in \mathbb{R}$, and $Z \sim N(0, \delta)$,

$$\int_{-a}^a e^{bs} f_Z(s) ds = e^{b^2 \delta / 2} \left[\Phi\left(\frac{a - b\delta}{\sqrt{\delta}}\right) - \Phi\left(\frac{-a - b\delta}{\sqrt{\delta}}\right) \right].$$

Hence we can compute

$$\begin{aligned} & \int_{-M\delta}^{M\delta} (e^{-2A_k} + e^{-2B_k}) f_Z(s) ds \\ &= \int_{-M\delta}^{M\delta} e^{-2E_k^2 \delta + 2E_k s} f_Z(s) ds + \int_{-M\delta}^{M\delta} e^{-2E_k^2 \delta - 2E_k s} f_Z(s) ds \\ &= e^{-2E_k^2 \delta} e^{(2E_k)^2 \delta / 2} \left[\Phi\left(\frac{M\delta - 2E_k \delta}{\sqrt{\delta}}\right) - \Phi\left(\frac{-M\delta - 2E_k \delta}{\sqrt{\delta}}\right) \right] \\ & \quad + e^{-2E_k^2 \delta} e^{(-2E_k)^2 \delta / 2} \left[\Phi\left(\frac{M\delta + 2E_k \delta}{\sqrt{\delta}}\right) - \Phi\left(\frac{-M\delta + 2E_k \delta}{\sqrt{\delta}}\right) \right] \\ &= [\Phi(-(4k-3)M\sqrt{\delta}) - \Phi(-(4k-1)M\sqrt{\delta}) + \Phi((4k-1)M\sqrt{\delta}) - \Phi((4k-3)M\sqrt{\delta})] \end{aligned}$$

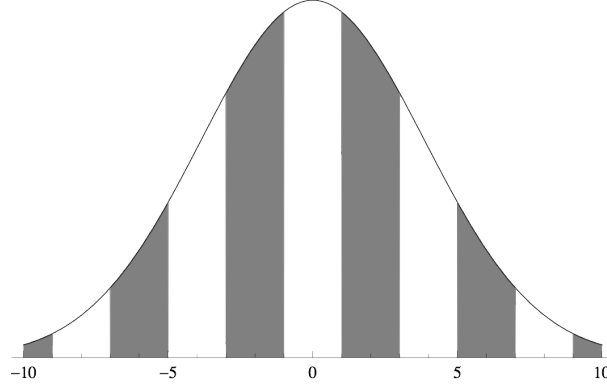


FIG 1. Twice the shaded region gives the required probability. The scale is in terms of $M\sqrt{\delta}$.

$$= 2[\Phi((4k-1)M\sqrt{\delta}) - \Phi((4k-3)M\sqrt{\delta})].$$

Similarly,

$$\int_{-M\delta}^{M\delta} (e^{-2C_k} + e^{-2D_k}) f_Z(s) ds = 2[\Phi((4k+1)M\sqrt{\delta}) - \Phi((4k-1)M\sqrt{\delta})]$$

Therefore

$$\begin{aligned} & P\left(\sup_{t \geq \delta} \frac{|W(t)|}{t} > M\right) \\ &= 1 - \int_{s=-M\delta}^{M\delta} \left[1 - \sum_{k=1}^{\infty} (e^{-2A_k} + e^{-2B_k} - e^{-2C_k} - e^{-2D_k})\right] f_Z(s) ds. \\ &= 2(1 - \Phi(M\sqrt{\delta})) + 2 \sum_{k=1}^{\infty} [2\Phi((4k-1)M\sqrt{\delta}) - \Phi((4k-3)M\sqrt{\delta}) - \Phi((4k+1)M\sqrt{\delta})] \\ &= 4 \sum_{k=1}^{\infty} [\Phi((4k-1)M\sqrt{\delta}) - \Phi((4k-3)M\sqrt{\delta})]. \end{aligned}$$

□

Remark 5.2. Figure 1 depicts the kind of integral we are computing in order to calculate the required probability. It is clear the number is very close to 1 if $M\sqrt{\delta}$ is small.

We can approximate the above infinite sum by a finite sum which can be optimized depending on our choice of M and δ . If we fix a small δ , i.e., ϵ now we can numerically approximate the quantiles of the distribution of $\sup_{t \geq \delta} \frac{|W(t)|}{t}$ using Proposition 5.1. This is what we use for creating confidence bands for the QQ plots in the examples in Section 6. Simulation suggests that considering the first 15 terms of the infinite sum is enough to give us approximations correct up to the first 6 decimal places.

5.1.1. Confidence band for QQ plots

We will provide a confidence band for a truncated version of the QQ plot defined by \mathcal{Q}_n in (3.2) where we truncate near zero for reasons explained earlier. For some truncation level $0 < \epsilon < 1$, ideally close to zero,

let

$$\mathcal{Q}_n^\epsilon := \left\{ \left(-\log \frac{j}{k}, \log \frac{X_{(j)}}{X_{(k)}} \right) : 1 \leq j \leq k \text{ and } \frac{j}{k} \geq \epsilon \right\}, \quad k < n, \quad (5.4)$$

Clearly under this truncation, Theorem 3.3 would imply that $\mathcal{Q}\mathcal{N}_n^\epsilon \Rightarrow \mathcal{Q}\mathcal{N}^\epsilon$, where $\mathcal{Q}\mathcal{N}_n^\epsilon$ and $\mathcal{Q}\mathcal{N}^\epsilon$ are the truncations of $\mathcal{Q}\mathcal{N}_n$ and $\mathcal{Q}\mathcal{N}$ respectively. Now for $0 < \alpha < 1$, the $(1 - \alpha)100\%$ confidence band for the limit QQ plot \mathcal{Q}^ϵ (\mathcal{Q} in (3.3) restricted to $t \in (\epsilon, 1)$) is:

$$\mathcal{Q}_n^\epsilon + \left\{ (0, y) : y \in \left(-\frac{c_{\alpha/2, \delta}}{\sqrt{k}}, \frac{c_{\alpha/2, \delta}}{\sqrt{k}} \right) \right\}, \quad (5.5)$$

where $c_{\alpha, \delta}$ is the $(1 - \alpha)$ -th quantile of $\sup_{t \geq \delta} \frac{W(t)}{t}$ with $\delta = \frac{\epsilon}{1 - \epsilon}$. One can calculate $c_{\alpha, \delta}$ using Proposition 5.1 now.

An important point to note here is that we are suggesting use the weak limit of the QQ plot to obtain the confidence band. In practice even if we have a large data set it will always be finite. A natural question that arises here is what is the rate of convergence in these cases. We do not have the answer at the moment but all the simulation study that we have done strongly suggest that this method works well.

5.2. ME plots

Under the assumption that the distributional tail $\bar{F} \in RV_{-1/\xi}$ with $\xi > 0$ we have observed in Section 4 that the weak limits we obtain for ME plots are different for the three different cases we considered.

In the case $0 < \xi < 1/2$ (limit given in Theorem 4.3) where F has a second moment we obtain a Gaussian limit, in terms of Brownian Bridges and finite integrals of Brownian Bridges. In order to convert this result to obtain confidence bounds we need to compute boundary crossing probabilities for these Gaussian processes. Analytical solution for these probabilities are available for linear boundaries (Doob (1949)), piecewise linear boundaries (Pötzelberger et al., 2001) in case of Brownian motion on $[0, \infty)$. Non-linear boundaries are usually approximated using piecewise linear boundaries. For applying Theorem 4.3 we require such results for non-linear boundaries on functionals of Brownian Bridges which is not readily available. Hence we resort to Monte Carlo simulation which is described in detail in Section 6.

For $1/2 < \xi < 1$ we know that the second moment of F does not exist but the mean exists. Theorem 4.6 provides the limit here where we have a functional of Brownian Bridge on the first component and a Stable distribution on the second component. One feature here is that the normalization required to get the limit depends on $b(n)$ and $b(n/k)$, which in turn depends on the distribution function F and is hence unknown. These can be estimated respectively with $X_{(1)}$ and $X_{(k)}$ in practice. Although to replace them theoretically we would need to know the joint behavior of sum of top k order statistics with the maxima and $k - th$ maxima of the distribution. Results akin to (Resnick, 1986, Section 4) can come handy here. See Darling (1952), Chow and Teugels (1978). Using Theorem 5.3 in Darling (1952) we can show that under the assumptions of Theorem 4.6

$$\begin{aligned} \widetilde{\mathcal{M}\mathcal{N}}_n &:= \left\{ \left(\left(\frac{i}{k} \right)^{-\xi}, \frac{\xi}{1 - \xi} \left(\frac{i}{k} \right)^{-\xi} \right) \right. \\ &\quad \left. + \left(\sqrt{k} \left(\frac{X_{(i)}}{X_{(k)}} - \left(\frac{i}{k} \right)^{-\xi} \right), \frac{kX_{(k)}}{X_{(1)}} \left(\frac{\hat{M}(X_{(i)})}{X_{(k)}} - \frac{\xi}{1 - \xi} \left(\frac{i}{k} \right)^{-\xi} \right) \right) : i = 2, \dots, k \right\} \\ &\Rightarrow \widetilde{\mathcal{M}\mathcal{N}} := \left\{ \left(t^{-\xi} + \xi t^{-(1+\xi)} B(t), \frac{\xi}{1 - \xi} t^{-\xi} + t^{-1} \tilde{S}_{1/\xi} \right), 0 < t \leq 1 \right\} \quad \text{in } \mathcal{F}, \end{aligned} \quad (5.6)$$

where $\tilde{S}_{1/\xi}$ is independent of $B(t)$ and its characteristic function is of the form

$$E[e^{i\lambda \tilde{S}_{1/\xi}}] = e^{i\lambda} \left(1 + \frac{i\lambda}{1 - \xi} - \frac{1}{\xi} \int_0^1 (e^{it\lambda} - 1 - it\lambda) t^{-1-1/\xi} dt \right)^{-1}. \quad (5.7)$$

We can again use simulation to obtain the quantiles of this distribution.

For $\xi \geq 1$, F need not have a finite mean and the ME plot does not have any nontrivial, nonrandom limit. We have shown weak limits in Theorem 4.9. Therefore, it is not sensible to obtain confidence bands in this case and hence is not pursued.

5.2.1. Confidence band for ME plots

We need to truncate the ME plot near infinity in this case since the weak limits we obtain (Theorems 4.3 and 4.6) blow up there (relates to t near 0 in the limit \mathcal{MN}_n). If \mathcal{S}_n , as defined in (4.1) denotes the ME plot for a sample of size n (with $k < n$ top order statistics under consideration), then for $0 < \epsilon < 1$ small, let \mathcal{S}_n^ϵ denote \mathcal{S}_n restricted to $1 \leq j \leq k$ with $j/k \geq \epsilon$. If $0 < \xi < 1/2$ then using Theorem 4.3 we can give the $(1 - \alpha)100\%$ confidence band for \mathcal{S}^ϵ (\mathcal{S} in (4.1) restricted to $t \in (\epsilon, 1)$) as

$$\mathcal{S}_n^\epsilon + \left\{ (x, y) : x \in \left(-\frac{c_{\alpha/2, \epsilon}}{\sqrt{k}}, \frac{c_{\alpha/2, \epsilon}}{\sqrt{k}} \right), y \in \left(-\frac{d_{\alpha/2, \epsilon}}{\sqrt{k}}, \frac{d_{\alpha/2, \epsilon}}{\sqrt{k}} \right) \right\}, \quad (5.8)$$

where

$$c_{\alpha, \epsilon} = (1 - \alpha)\text{-th quantile of } \sup_{\epsilon \leq t \leq 1} \xi t^{-(1+\xi)} B(t), \quad (5.9)$$

$$d_{\alpha, \epsilon} = (1 - \alpha)\text{-th quantile of } \sup_{\epsilon \leq t \leq 1} \xi t^{-1} \int_0^t y^{-(1+\xi)} B(y) dy.$$

If $1/2 < \xi < 1$ then we use Theorem 4.6 and its modified form in (5.6) to give the $(1 - \alpha)100\%$ confidence band for \mathcal{S}^ϵ as

$$\bigcup_{\epsilon k \leq j \leq k} \left\{ \left(\frac{X_{(j)}}{X_{(k)}}, \frac{\hat{M}(X_{(j)})}{X_{(k)}} \right) + \left(-\frac{c_{\alpha/2, \epsilon}}{\sqrt{k}}, \frac{c_{\alpha/2, \epsilon}}{\sqrt{k}} \right) \times \left(\frac{X_{(1)} d_{(1-\alpha)/2}}{j X_{(k)}}, \frac{X_{(1)} d_{\alpha/2}}{j X_{(k)}} \right) \right\}, \quad (5.10)$$

where

$$d_\alpha = (1 - \alpha)\text{-th quantile of } \tilde{S}_{1/\xi} \text{ defined in (5.7).}$$

The above quantiles are calculated by Monte Carlo for the simulation we report in Section 6.1.2.

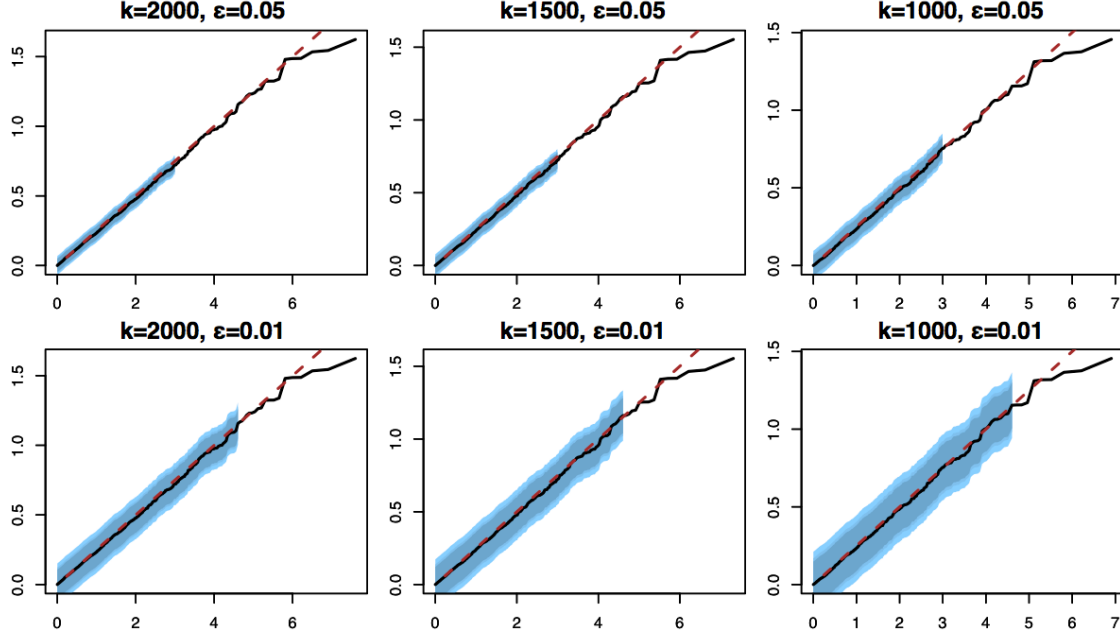
Remark 5.3. Throughout the literature of extreme value theory people always consider the top k order statistics where $k = k_n \rightarrow \infty$ and $k/n \rightarrow 0$. The idea is that as the size of data increases we concentrate more on the right end point of the underlying distribution. In practice though when you have a data with a fixed size, albeit large, it is difficult to decide on which value of k to choose. The popular solution is to try out different values of k which have been exemplified in the Hill plot and the Pickands plot.

In order to obtain confidence band for the QQ plot and the ME plot along with the problem of choosing k we also have to choose ϵ . The choice of ϵ should be such that, for the purpose of drawing any inference, we leave out the region where data is sparse. In practice we have to try out different values of ϵ depending on the size of the data and the choice of k .

6. QQ Plot and ME Plot in Practice

6.1. Simulation

We do a simulation study using the software R to check how well this method of obtaining confidence bounds for the QQ plot and the ME plot work.

FIG 2. QQ Plot for 50000 i.i.d. Pareto random variables with $\xi = 0.25$.

6.1.1. QQ plots

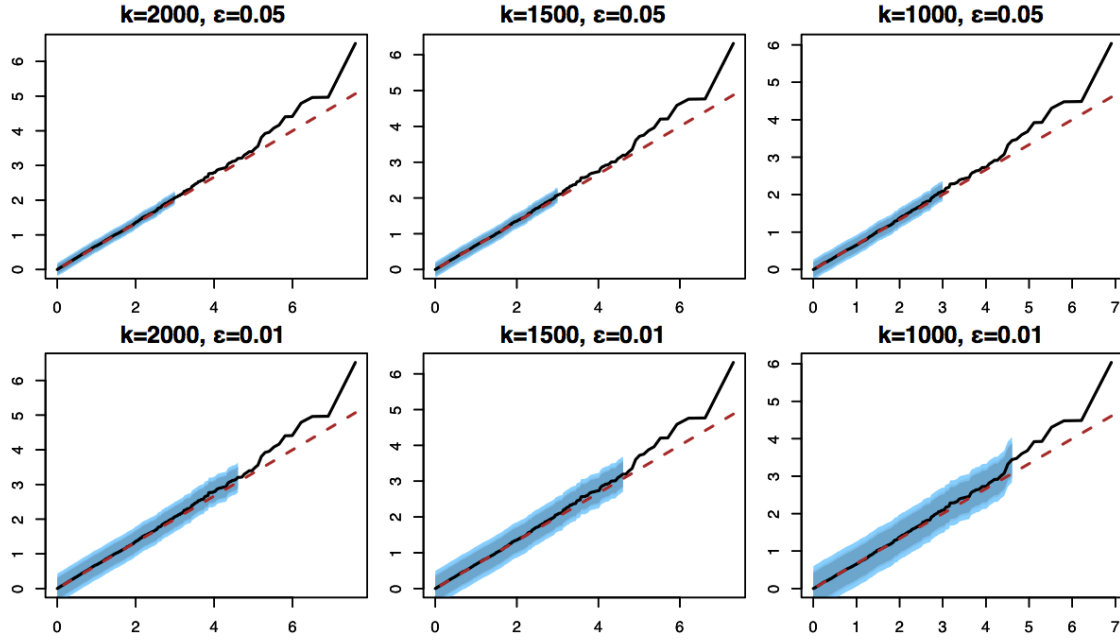
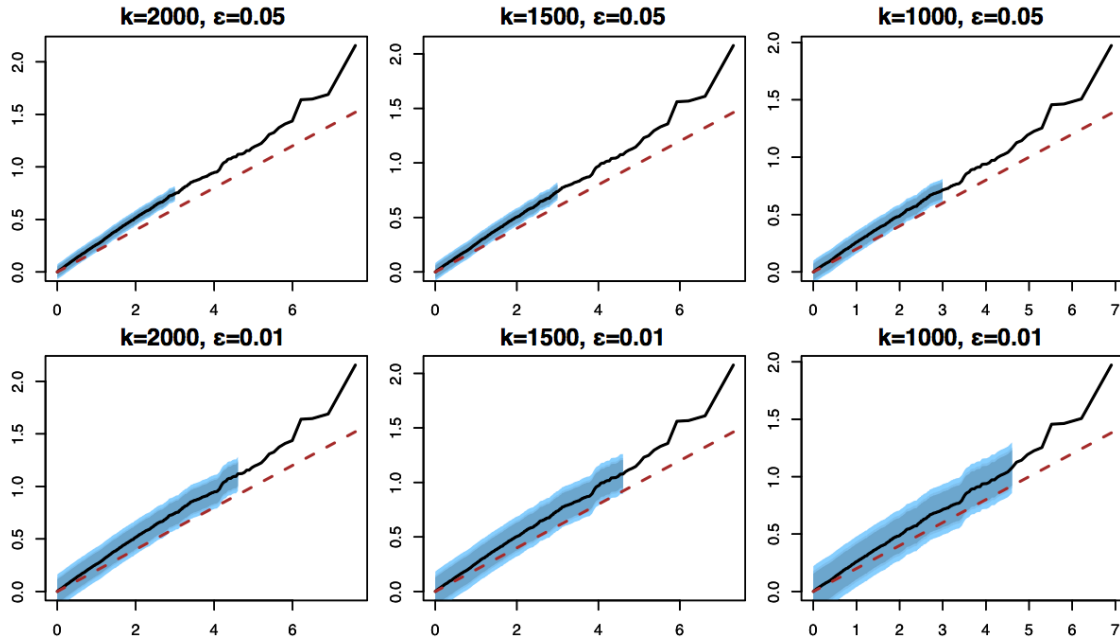
We begin with a simple exercise for Pareto distribution with $\xi = 0.25$ ($\bar{F}(x) = x^{-4}, x \geq 1$). We simulate a sample of size $n = 50000$ from this distribution and look at the QQ plot for extremes as defined in (3.2), see Figure 2. The black line denotes the plot Q_n and the brown dotted line denotes the true line Q . We know that Q_n converges to Q and as we see in the plot the two lines are close except for the top right corner of the plots which correspond to the very large order statistics. We choose 3 different values for k : 2000, 1500 and 1000 which are large in absolute terms but small compared to the sample size n .

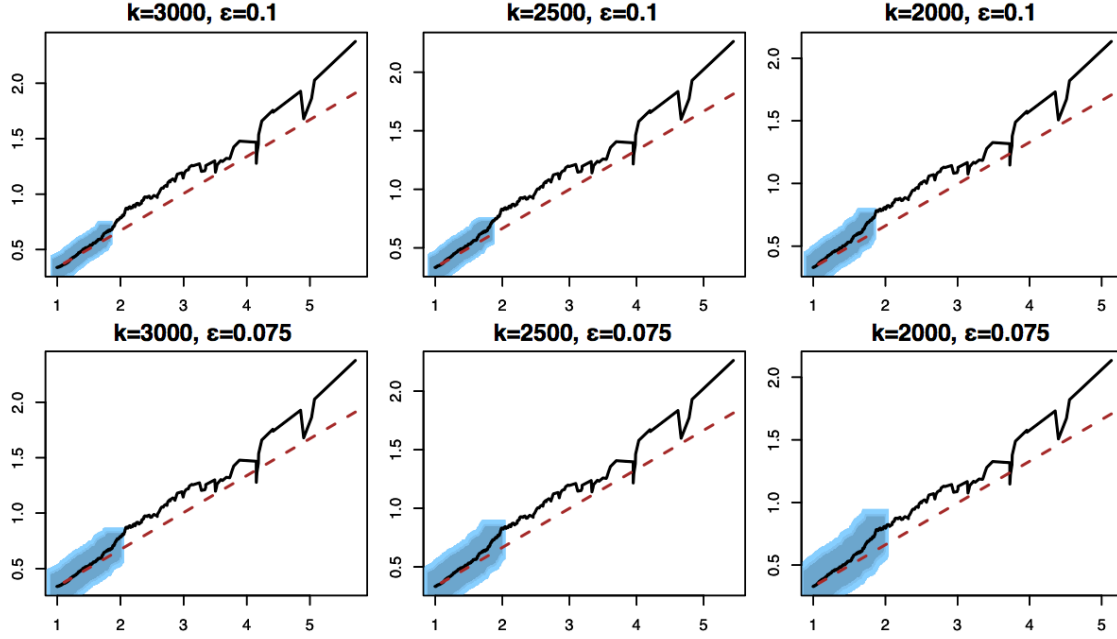
Following the discussion in Section 5 we know that the variance of the limiting distribution blows up as we move towards the extreme order statistics (towards the top-right corner) in the plot. So while obtaining a confidence bound we truncate at $\lfloor \epsilon k \rfloor$ -th order statistic for $\epsilon = 0.05$ and 0.01 . The confidence bounds are obtained for the 6 cases. The three shades of the colored band signifies the 99%, 95% and the 90% confidence band for the plot. As is evident in Figure 2 the true line lies within the bound in all the cases. It is also notable that the width of the confidence band increases as k and ϵ decreases.

We use the result in 5.1 to compute the width of confidence band. Also note that the limit depends on the value of ξ and while obtaining the width of the band we replace ξ by its Hill estimate (Resnick, 2007, p.74). We could use any consistent estimator of ξ and the choice of the estimator does not seem to be important as far as the simulation study is concerned. It is well known that estimating the parameter ξ can often be extremely tricky, see ‘‘Hill-Horror plots’’ in (Resnick, 2007, p.87). But as far as obtaining confidence bounds is concerned we can get past that by using a conservative estimate of ξ , i.e. a value which we strongly believe is not less than the true value of ξ .

Next we do a similar study for a right-skewed stable distribution with $\xi = 2/3$ ($\alpha = 1.5$) and mean 0. We use the same values of n, k and ϵ . The result is given in Figure 3. Here also we see that the method works well and the confidence band contains the true line in all the 6 cases.

We also try a non-standard distribution for which $\bar{F}^{-1}(x) = x^{-1/5}(1 - 10^{-1} \ln x), 0 < x \leq 1$. This means

FIG 3. QQ Plot for 50000 i.i.d. right-skewed stable random variables with $\xi = 2/3$.FIG 4. QQ Plot for 50000 i.i.d. random variables with the distribution described in (6.1) ($\xi = 0.2$).

FIG 5. ME Plot for 50000 i.i.d. Pareto random variables with $\xi = 0.25$.

that $\bar{F} \in RV_{-4}$ and therefore $\xi = 0.5$. The exact form of \bar{F} is given by

$$\bar{F}(x) = \frac{1}{32} W(2xe^2)^5 x^{-5} \quad \text{for all } x \geq 1, \quad (6.1)$$

where W is the Lambert W function satisfying $W(x)e^{W(x)} = x$ for all $x > 0$. Observe that $W(x) \rightarrow \infty$ as $x \rightarrow \infty$ and $W(x) \leq \log(x)$ for $x > 1$. Furthermore,

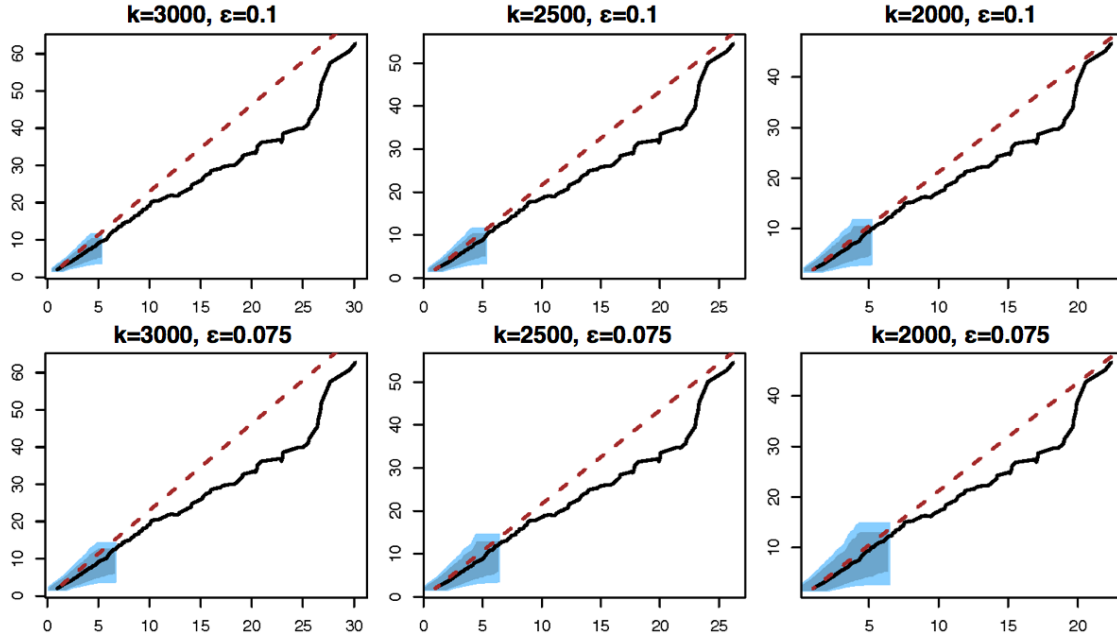
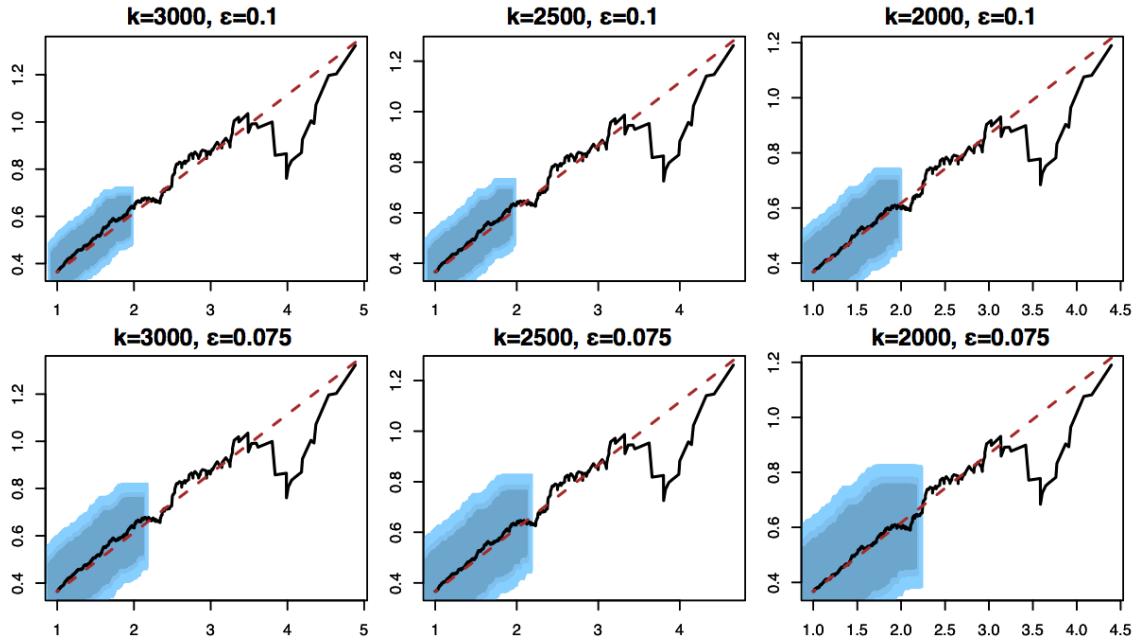
$$\frac{\log(x)}{W(x)} = 1 + \frac{\log W(x)}{W(x)} \rightarrow 1 \quad \text{as } x \rightarrow \infty,$$

and hence $W(x)$ is a slowly varying function. This is therefore an example where the slowly varying term contributes significantly to \bar{F} . That was not the case in the Pareto or the stable examples. The result of the simulation is shown in Figure 4. As expected, the choice of k plays an important role in this case and we see that the confidence band contains the true line when we choose $k = 1000$ and $\epsilon = 0.01$. Although not shown in Figure 4 the confidence bands perform better for smaller values of k .

6.1.2. ME plots

Figure 5 shows the ME plot obtained from a data simulated from the Pareto distribution with $\xi = 0.25$. The 6 plots correspond to different values of k (3000, 2500 and 2000) and ϵ (0.1 and 0.075). The black line is the observed ME plot and the brown dotted line denotes the limit in probability. Again, the three shades of the colored band signifies the 99%, 95% and the 90% confidence band for the plot. Note that the weak limit is a functional of the Brownian bridge and depends on ξ . We estimate ξ using the Hill estimator and obtain the bounds by simulating 10000 paths from the weak limit.

A striking feature in all these plots is that they are close to being linear near the bottom left corner and becomes quite erratic near top right corner. The reason behind this phenomenon is that the empirical ME function for high thresholds is the average of the excesses of a small number of upper order statistics.

FIG 6. ME Plot for 50000 i.i.d. right-skewed stable random variables with $\xi = 2/3$.FIG 7. ME Plot for 50000 i.i.d. random variables with the distribution described in (6.1) ($\xi = 0.2$).

When averaging over few numbers, there is high variability and therefore, this part of the plot appears very non-linear and is uninformative. Therefore, while obtaining confidence bands it is essential to leave out some of the extreme order statistics. We would also like to point out that without the confidence bands it would have been difficult to believe that these plots were obtained from a a distribution with tail index 0.25.

A simulation of ME plot for the right skewed stable distribution with $\xi = 2/3$ is shown in Figure 6. We use the band described in (5.10) and estimate the quantiles using simulation. In this case we only provide the 95% and the 90% confidence band. The 99% confidence band for the stable is very large and using that is not much helpful.

The next simulation is the ME plot for a sample from the distribution function described in (6.1) and the result is given in Figure 7. We use the same values for n, k and ϵ . We see that this method of getting confidence bands works well in these cases.

6.2. An example with a real data

We study a data set which contains Internet response sizes corresponding to user requests. The sizes are thresholded to be at least 100KB. The data set consists of 67287 observations and is part of a bigger set collected in April 2000 at the University of North Carolina at Chapel Hill.

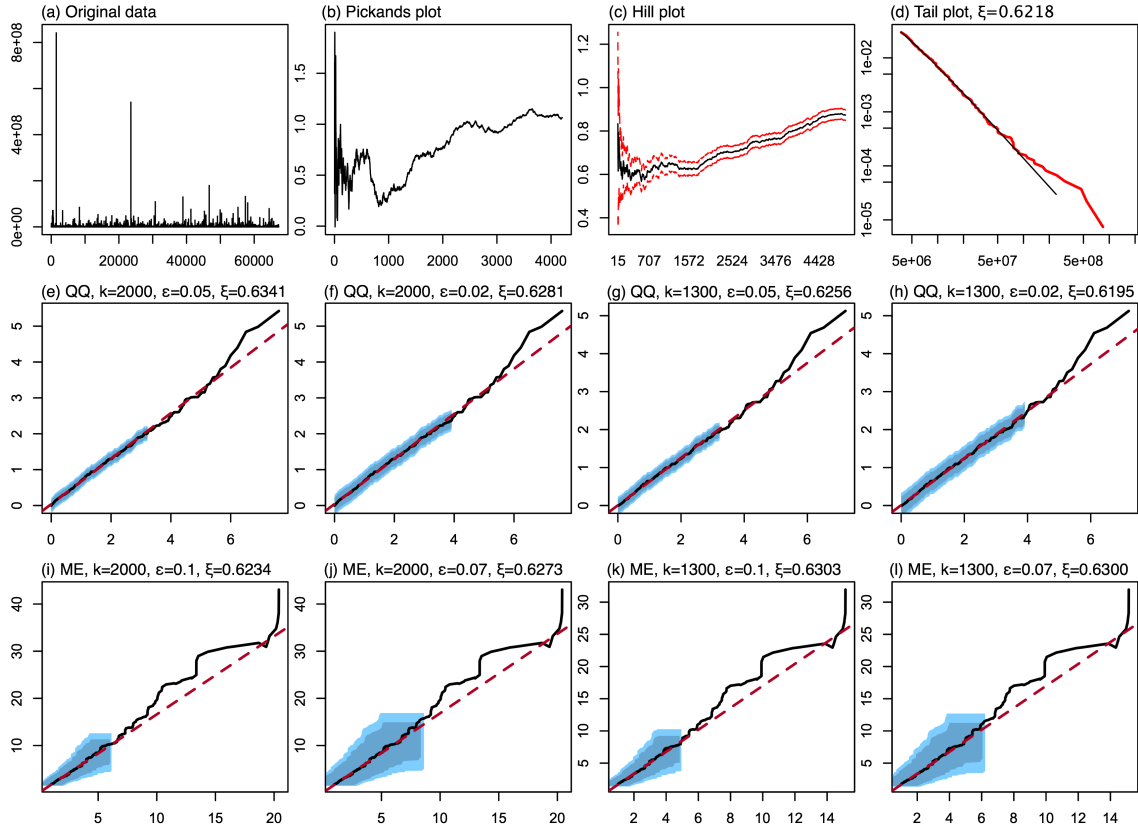


FIG 8. Analysis of the internet response size data.

It is often stated that file size data typically exhibits heavy-tails and we observe that is indeed the case here. Figure 8 shows various plots from this data set. The sample variance is of the order of 10^{13} which

suggests that the variance is possibly infinite for the underlying distribution (denote by F). This would imply that if \bar{F} is regularly varying for some ξ then we must have $\xi \geq 1/2$. This is suggested by both the Pickands plot and the Hill plot (Figure 8(b) and (c) respectively). The Hill plot is always above $1/2$ and the Pickands is above $1/2$ for most of the range. But it is difficult to get an estimate of ξ using these two tools since both the plots are highly fluctuating and hence inconclusive. We fit a GPD model with the top 2000 order statistics using the command ‘fit.GPD’ in the library ‘QRMLib’. It gives an estimate 0.6218 of ξ and Figure 8(d) plots the estimated \bar{F} in the log-log scale along with the fitted line.

We try the QQ plot with data set for $k = 2000$ and 1300 (top 3% and 2% order statistics approximately) and with $\epsilon = 0.05$ and 0.02 . The plots give an estimate of around 0.62-0.63 of ξ . The plots are shown in Figure 8(e)-(h). The ME plots for $k = 2000, 1300$ and $\epsilon = 0.1, 0.07$ are shown in Figure 8(i)-(l) and they also suggest a similar range for ξ .

We observe that in this example the different methods of understanding the tail behavior of a data work very well and all of them are in agreement about the value of ξ . This is not true in many situations and then it is hard to judge which method one should trust. In those cases it is important to have some more knowledge about the system from which the data was collected and often that helps in the understanding of the data.

7. Conclusion

Plotting techniques have always been popular as diagnostic tools for goodness-of-fit of observed data and we believe they will remain so because of their visual and intuitive appeal. In this paper we have concentrated on two such tools used extensively in the extreme-value literature. A weak law of large numbers have been shown previously for both the QQ plots (Das and Resnick, 2008) and ME plots (Ghosh and Resnick, 2010) considering them as random elements in an appropriate topology. Our contribution in this paper has been to provide distributional limits for them. In the case of QQ plots we have also provided an explicit expression for confidence bounds (with a truncation to avoid the confidence bounds from blowing up) by using these distributional results. In the case of ME plots we have obtained distributional limits in the cases $0 < \xi < 1/2$, $1/2 < \xi < 1$ and $\xi \geq 1$ separately where the underlying distribution F is assumed to be regularly varying with index $-1/\xi$. The case $\xi = 1/2$ is still open. We have produced confidence bounds for the ME plots in these cases by Monte Carlo simulation, as explicit expressions for these quantities are not easy to calculate. The explicit expressions would involve boundary crossing probabilities for a Brownian Bridge with non-linear boundaries. Boundary crossing probabilities for Brownian motion can be approximated using piecewise linear boundaries (Pötzelberger et al., 2001) but we do not know of a nice approximation for the Brownian Bridge case, hence we resort to simulation. We have illustrated the confidence bounds in both the cases of QQ plots and ME plots with simulated and real data examples in Section 6. The importance of the confidence bounds can be understood very clearly from Figure 5. Here we have a simulated data set of 50000 points from a Pareto distribution with parameter $\xi = 0.25$. Just looking at ME plot it is not at all obvious that this is a heavy-tailed data whereas when the confidence bounds with the ϵ -truncation are drawn, the straight line with slope $\xi = 0.25$ remains inside the bounds indicating the true nature of the data.

Since we are using the limiting distribution to obtain the confidence bounds it is natural to ask what the rate of convergence is. We have observed that this method works well in the simulation studies that we have done but we have not answered this theoretically. This is currently a work in progress.

A standing assumption in the results we proved in this paper is that the random variables X_n are iid. We believe that it is possible to obtain similar results under a more general assumption of stationarity and mixing, cf. Rootzén (2009). We intend to look into this further.

We should also note here that often practitioners use the median-excess plot with the implied meaning when $\xi > 1$, i.e., the mean for the distribution does not exist (Embrechts et al., 1997), but we have not ventured into this kind of plotting tool. We have also not looked into other kinds of plots used in extremes like the Stărică plot (Stărică, 1999) to determine the right k number of upper order statistics, or the Gertensgarbe and Werner plot (Gertensgarbe and Werner, 1989) for determining thresholds over which a data may be assumed to be extreme-valued or the more popular Hill plot, Pickands plot (Resnick, 2007) to detect the

right value of the extreme-value parameter. Obtaining results in the same spirit as this paper for these other varieties of plots are a part of intended future research.

References

- P. Billingsley. *Convergence of Probability Measures*. John Wiley & Sons Inc., New York, 1968.
- N.H. Bingham, C.M. Goldie, and J.L. Teugels. *Regular Variation*. Cambridge University Press, 1987.
- T.L. Chow and J.L. Teugels. The sum and the maximum of iid random variables. *Proceedings of the Second Prague Symposium on Asymptotic Statistics*.
- S. Csörgö, L. Horváth and D.M. Mason. The asymptotic distribution of extreme sums *The Annals of Probability*, 19(2):783–811, 1991.
- S. Csörgö, L. Horváth and D.M. Mason. What portion of the sample makes a partial sum asymptotically stable or normal? *Probability Theory and Related Fields*, 72(1):1–16, 1986.
- M. Csörgö and D.M. Mason. On the asymptotic distribution of weighted uniform empirical and quantile processes in the middle and on the tails. *Stochastic Processes and their Applications*, 21:119–132, 1985.
- D.A. Darling. The influence of the maximum term in the addition of independent random variables. *Transactions of the American Mathematical Society*, 73(1):95–107, 1952.
- B. Das and S.I. Resnick. QQ plots, random sets and data from a heavy tailed distribution. *Stochastic Models*, 24(1):103–132, 2008. ISSN 1532-6349.
- B. D’Auria and S.I. Resnick. Data network models of burstiness. *Advances in Applied Probability*, 38(2):373–404, June 2006.
- A.C. Davison and R.L. Smith. Models for exceedances over high thresholds. (with discussion). *Journal of the Royal Statistical Society, Series B*, 52:393–442, 1990.
- L. de Haan. *On Regular Variation and Its Application to the Weak Convergence of Sample Extremes*. Mathematisch Centrum Amsterdam, 1970.
- L. de Haan and A. Ferreira. *Extreme Value Theory: An Introduction*. Springer-Verlag, New York, 2006.
- L. de Haan and L. Peng. Comparison of tail index estimators. *Statistica Neerlandica*, 52:60–70, 1998.
- L. de Haan and U. Stadtmueller. Generalized regular variation of second order. *Journal of Australian Mathematical Society, Series A*, 61(3):381–395, 1996.
- J.L. Doob. Heuristic approach to the Kolmogorov-Smirnov theorems. *Annals of Mathematical Statistics*, 20:393–403, 1949. ISSN 0003-4851.
- P. Embrechts, C. Kluppelberg, and T. Mikosch. *Modelling Extreme Events for Insurance and Finance*. Springer-Verlag, Berlin, 1997.
- J.L. Geluk and L. de Haan. *Regular Variation, Extensions and Tauberian Theorems*, volume 40 of *CWI Tract*. Stichting Mathematisch Centrum, Centrum voor Wiskunde en Informatica, Amsterdam, 1987. ISBN 90-6196-324-9.
- F.W. Gertensgarbe and P.C. Werner. A Method for the Statistical Definition of Extreme-Value Regions and Their Application to Meteorological Time Series. *Zeitschrift für Meteorologie* 39:224–226.
- S. Ghosh and S.I. Resnick. A discussion on mean excess plots. *Stochastic Processes and their Applications*, 120:1492–1517, 2010.
- R. Hogg and S. Klugman. *Loss Distributions*. Wiley, New York, 1984.
- R.W. Katz, M.B. Parlange, and P. Naveau. Statistics of extremes in hydrology. *Advances in Water Resources*, 25:1287–1304, 2002.
- M. Kratz and S.I. Resnick. The QQ-estimator and heavy tails. *Stochastic Models*, 12:699–724, 1996.
- G. Matheron. *Random Sets and Integral Geometry*. John Wiley & Sons, New York-London-Sydney, 1975. With a foreword by G.S. Watson, Wiley Series in Probability and Mathematical Statistics.
- K. Maulik, S.I. Resnick, and H. Rootzén. Asymptotic independence and a network traffic model. *Journal of Applied Probability*, 39(4):671–699, 2002. ISSN 0021-9002.
- A.J. McNeil, R. Frey, and P. Embrechts. *Quantitative Risk Management: Concepts, techniques and tools*. Princeton Series in Finance. Princeton University Press, Princeton, NJ, 2005. ISBN 0-691-12255-5.

- I. Molchanov. *Theory of Random Sets*. Probability and its Applications (New York). Springer-Verlag London Ltd., London, 2005. ISBN 978-185223-892-3; 1-85233-892-X.
- B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer, 2003. ISBN 978-354004-758-2; 3-54004-758-1.
- L. Peng. Second Order Condition and Extreme Value Theory. *PhD Thesis*. Tinbergen Institute, Tinbergen Institute Research Series 178. Erasmus University, Rotterdam, 1998.
- K. Pötzelberger, and L. Wang. Boundary crossing probability for Brownian motion. *Journal of Applied Probability*, 38:152-164, 2001.
- S.I. Resnick. Point processes, regular variation and weak convergence. *Advances in Applied Probability*, 18: 66-138, 1986.
- S.I. Resnick. *Heavy Tail Phenomena: Probabilistic and Statistical Modeling*. Springer Series in Operations Research and Financial Engineering. Springer-Verlag, New York, 2007. ISBN: 0-387-24272-4.
- S.I. Resnick. *Extreme Values, Regular Variation and Point Processes*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2008. ISBN 978-0-387-75952-4. Reprint of the 1987 original.
- H. Rootzén. Weak convergence of the tail empirical process for dependent sequences. *Stochastic Processes and their Applications*, 119:468–490, 2009.
- H.J. Rossberg. Über das asymptotische Verhalten der Rand-und Zentralglieder einer Variationsreihe (II). *Publicationes Mathematicae Debrecen*, 14, 83–90, 1967.
- E. Seneta. *Regularly Varying Functions*. Springer-Verlag, New York, 1976. Lecture Notes in Mathematics, 508.
- G.R. Shorack and J.A. Wellner. *Empirical Processes with Applications to Statistics*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. ISBN 0-471-86725-X.
- R.L. Smith. Statistics of extremes, with applications in environment, insurance and finance. In B. Finkenstadt and H. Rootzén, editors, *SemStat: Seminaire Europeen de Statistique, Exteme Values in Finance, Telecommunications, and the Environment*, pages 1–78. Chapman-Hall, London, 2003.
- C. Stărică. Multivariate extremes for models with constant conditional correlations. *Journal of Empirical Finance*, 6:515–553, 1999.
- G.L. Yang. Estimation of a biometric function. *Annals of Statistics*, 6(1):112–116, 1978.
- W. Vervaat. Random upper semicontinuous functions and extremal processes. In W. Vervaat and H. Holwerda, editors, *Probability and Lattices*, volume 110 of *CWI Tract*, pages 1–56. Math. Centrum Wisk. Inform., Amsterdam, 1997.